TECHNICAL REPORT

Feature Extraction and Classification of Magnetic and EMI Data, Camp Beale, CA

ESTCP Project MR-201004

MAY 2012

Leonard Pasion
Sky Research, Inc.

ESTCP

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAY 2012	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Feature Extraction and Classification of Magnetic and EMI Data, Camp Beale, CA	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Sky Research, Inc.**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
ESTCP Project MR-201004

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**117**	

Executive Summary

The demonstration described in this report was conducted at the Former Camp Beale, California, under project ESTCP MR-201004 "Practical Strategies for UXO Discrimination." It was performed under the umbrella of the ESTCP Live-Site Classification Study Program. The objective of the MR-201004 project is to demonstrate the application of feature extraction and statistical classification to the problem of UXO discrimination. At the Camp Beale site, the objective was to discriminate targets of interest (TOI) (including 37 mm, 60 mm, 81 mm targets, 105 mm and a small industry standard object (ISO)) from non-hazardous shrapnel, range and cultural debris. A number of fuses and fuse parts that were initially considered as a possible TOI was later found to be non-TOI, and labeled as clutter. In this report, we describe the performance of classification techniques that utilized full coverage, dynamically acquired, survey data acquired with a Geonics EM61 cart and static, cued interrogation style data acquired with advanced electromagnetic induction (EMI) sensors. Analysts from Sky Research and UBC-GIF processed (1) MetalMapper data acquired in a portion of the site amenable to vehicular towed systems (i.e. the "Open Area"), and (2) man portable sensor data collected by the TEMTADS 2x2, BUD, and MPV in a treed section of the site (i.e. the "Portable Area"). A small overlap of the Open and Portable Area allowed for a direct comparison of the MetalMapper and portable EMI sensors.

The classification techniques applied to the Camp Beale data use dipole model-based features extracted from observed data around each anomaly. One source and two source dipole inversions were carried out for all data sets, with the exception of the MPV data for which only a single source inversion was carried out. The feature sets used for classification were (1) the dipole polarizability tensor for the advanced EMI sensors, and (2) data magnitude and polarizability decay rate for the Geonics EM61 data. From the extracted feature vectors the following prioritized dig-lists were created: (i) EM61 cart data ranked by polarizability time-decay and data amplitude; (ii) MetalMapper statistical classification using a Support Vector Machine (SVM) applied to polarizabilities; (iii) MetalMapper classification using a library matching method (iv) TEMTADS 2x2 statistical classification using a SVM applied to polarizabilities; (v) BUD classification using library matching; (vi) MPV classification using a SVM applied to polarizabilities estimated using a single source inversion; and (vii) MPV classification using a library matching method using polarizabilities estimated using a single source inversion. A version of the MetalMapper diglists that were prescreened using the EM61 data was also submitted. All model fits and discrimination analysis were performed using the Sky classification software suite (UXOLab) that was jointly developed by UBC-GIF and Sky Research.

In order to preserve the integrity of the blind-test, and to provide a comparison between the different processing methods, a unique analyst was assigned to process each data set/processing approach combination. Seven different analysts from Sky Research and UBC-GIF processed the data sets. In this way, ground truth obtained as part of processing data with a particular method would not be shared with different classification approach. For each dig-list, an objective classification method was applied to the features to construct the dig-list. Analysts were instructed to avoid "expert input" when constructing the dig-list. The dig order was determined automatically, and analysts did not manually change the position of any anomaly within a list based on visual inspection of the polarizabilities (i.e. manual labeling of anomalies as TOI or non-TOI).

Geonics EM61 cart data were collected over both the Open Area and the Portable Area. The approach with the EM61 data was to apply an automated method (i.e. minimal QC for inversions or data) and to

apply a classifier that used data amplitude and the decay rate of the estimated polarizability tensor. The data amplitude/decay classifier did surprisingly well identifying larger TOI (i.e. 105 mm and 81 mm). For the Portable Area, the analysis required approximately 30% of clutter to be excavated in order to find all TOI. For the Open Area, approximately 50% of clutter would be required to be excavated in order to find all TOI. For both the Open and Portable Area data, we did not have confidence in the ability to classify the smaller TOI (i.e. small ISO and 37 mm). Therefore we assigned very conservative stop dig points for both data sets (89.0% and 88.7% scrap dug at the operating point for the Portable and Open Area, respectively).

MetalMapper data were acquired by Parsons and CH2M HILL in the Open Area, with a different analyst processing each data set. Application of a library classification method to the CH2M HILL data resulted in 99.2% of TOI identified (due to one missed TOI) and 27.4% scrap dug at the stop dig point. The SVM based classification method resulted in two missed TOI, such that 98.5% of TOI were identified and 38.9% of scrap were dug at the stop dig point. The prescreened versions of the CH2M HILL MetalMapper diglists did not miss any TOI (100% TOI identified at the stop-dig point). Application of a library classification method to the Parsons data resulted in 98.5% of TOI identified (due to two missed TOI) and 16.3% scrap dug at the stop-dig point. The SVM based classification method resulted in no missed TOI and 37.0% of scrap were dug at the stop dig point. The false alarm rate (FAR) for the SVM classifier was 20.2%, suggesting that the automated stop-dig point method was overly conservative. The analyst that applied the library method to the Parsons data chose an overly aggressive classification approach and missed two TOI whose data anomalies were unable to constrain the secondary polarizabilities. The two missed TOI would have been identified if a library match using only the primary polarizability was included in the classification approach. The SVM approach used two stage approach that first used all polarizabilities for classification, then used only the primary polarizability for classification. This two stage approach adopted by the SVM method was more conservative and was thus able to identify all TOI at the stop-dig point. The prescreened and non-prescreened versions of the Parsons MetalMapper dig-lists had similar performance.

TEMTADS 2x2, MPV and BUD data were collected in the Portable Area. For each of the portable sensors all TOI were found at the chosen operating point. The BUD and TEMTADS data were inverted for a single source and two source dipole models. The BUD library match method produced a FAR of 13.4%, with 26.2% of the scrap dug at the stop-dig point. The more conservative SVM classification approach for the BUD produced a FAR of 21.7%, with 37.9% of scrap dug at the operation point. Only an SVM dig-list was submitted for the TEMTADS data. The resulting FAR for the SVM list was 14.8%, with 19.5% of the scrap dug at the operating point. The similar FAR for of the TEMTADS and BUD dig-lists suggest that the approaches have similar ability to classification. The fewer number of scrap of the TEMTADS is a reflection of the aggressiveness of the analyst setting the stop dig point. The MPV data were only inverted for a single source. Using the polarizabilities recovered from the single source inversion produced a FAR of 25.0% and 36.6% for the SVM and Library Match, respectively. The SVM approach required 40.1% of the scrap to be dug at the operating point, and the Library Match approach required 37.9% of scrap to be dug at the operating.

There are a number of conclusions from applying dipole-based classification techniques to the Camp Beale demonstration. Firstly, the production quality EM61 data produced mediocre results due to the relative sizes of the scrap and the smallest UXO, and the inability to accurate constrain their size and shape. Secondly, the MetalMapper sensor deployed in the Open Area in a cued-interrogation mode

resulted in a significant reduction in the number of digs required to excavate all UXO. Diglists were submitted for both the Parsons and CH2M Hill data sets that resulted in having no TOI left in the ground at the operating point. A more conservative classification approach with the MetalMapper produced better results than aggressive techniques. Thirdly, all the portable sensors data sets produced dig-lists with no TOI left in the ground.

Table of Contents

List of Tables vi

List of Figures vi

Acronyms xii

1.0 INTRODUCTION 1

1.1 BACKGROUND 1
1.2 OBJECTIVE OF THE DEMONSTRATION 1
1.3 REGULATORY DRIVERS 3

2.0 TECHNOLOGY 4

2.1 TECHNOLOGY DESCRIPTION 4
2.1.1 Creation of a Map of Geophysical Sensor Data 5
2.1.2 Anomaly Selection and Feature Extraction 5
2.1.3 Classification of Anomalies 6
2.2 UXOLab Software 8
2.3 PREVIOUS TESTING OF THE TECHNOLOGY 9
2.4 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY 12

3.0 PERFORMANCE OBJECTIVES 13

3.1 OBJECTIVE: MAXIMIZE CORRECT CLASSIFICATION OF MUNITIONS 13
3.1.1 Metric 13
3.1.2 Data Requirements 13
3.1.3 Success Criteria 13
3.2 OBJECTIVE: MAXIMIZE CORRECT CLASSIFICATION OF NON-MUNITIONS 16
3.2.1 Metric 16
3.2.2 Data Requirements 16
3.2.3 Success Criteria 16
3.3 OBJECTIVE: SPECIFICATION OF NO-DIG THRESHOLD 16
3.3.1 Metric 16
3.3.2 Data Requirements 16
3.3.3 Success Criteria 16
3.4 OBJECTIVE: MINIMIZE NUMBER OF ANOMALIES THAT CANNOT BE ANALYZED 17
3.4.1 Metric 17
3.4.2 Data Requirements 17
3.4.3 Success Criteria 17
3.5 OBJECTIVE: CORRECT ESTIMATION OF TARGET PARAMETERS 17
3.5.1 Metric 17
3.5.2 Data Requirements 17
3.5.3 Success Criteria 17

4.0 SITE DESCRIPTION 18

4.1 SITE SELECTION 18
4.2 SITE HISTORY 18
4.3 SITE GEOLOGY 18
4.4 MUNITIONS CONTAMINATION 18

5.0 TEST DESIGN 19

6.0 DATA ANALYSIS AND RESULTS 20

6.1 Processing Portable Area Datasets 21
6.1.1 Geonics EM61 Mk2 Cart Results 21
6.1.2 HandHeld BUD Results 23
6.1.3 TEMTADS 2x2x3 Results 24
6.1.4 MPV Results 26
6.1.5 Portable Area Summary 28
6.2 Processing Open Area Datasets 30
6.2.1 Geonics EM61 Mk2 Cart Results 30
6.2.2 CH2M Hill MetalMapper Results 31
6.2.3 Parsons MetalMapper Results 34
6.2.4 Prescreened MetalMapper Results 36
6.2.5 Open Area Summary 37

7.0 MANAGEMENT AND STAFFING 39

8.0 REFERENCES 40

APPENDICES A-1

Appendix A: EM61 FEATURE EXTRACTION AND CLASSIFICATION A-1
Appendix B: METALMAPPER FEATURE EXTRACTION AND CLASSIFICATION B-1
Parsons MetalMapper static feature extraction and classification B-1
CH2M Hill MetalMapper Static Feature Extraction and Classification B-19
Appendix C: SUPPORT VECTOR MACHINE DIGLISTS C-1
Appendix D: PROCESSING AND CLASSIFICATION ANALYSIS OF BUD HAND-HELD DATA AT CAMP BEALE D-1
Appendix E: Target Location Error Analysis E-1
Appendix F: Points of Contact F-1

List of Tables

Table 0.1 Kernel functions commonly used with nonlinear support vector machines……8
Table 1. Previous Inversion/Classification Testing 10
Table 2. Performance Objectives for the Camp Beale Portable Area (fuzes as clutter) 14
Table 3. Performance Objectives for the Camp Beale Open Area (fuzes as clutter) 15
Table 4: Summary of performance for all sensors in the Portable area when fuzes are considered clutter. 29
Table 5: Summary of performance for all sensors in the Portable area when fuzes are considered TOI. .. 29
Table 6: Summary of performance for all sensors in the Open area when fuzes are considered clutter. ... 38
Table 7: Summary of performance for all sensors in the Open area when fuzes are considered TOI. 38
Table 8. Training data requests. Items highlighted in yellow are TOI. B-10
Table 9: Results of stage-1 dig list. PPV: Positive predictive value D-10
Table 10: Result of modified stage-1 dig list. PPV: Positive predictive value D-14
Table 11: Summary of target location error for MetalMapper datasets E-1

List of Figures

Figure 0.1: SVM formulation for constructing a decision boundary. The decision boundary bisects support planes bounding the classes……8

Figure 1: Full coverage EM61 data was collected and used to pick targets for cued investigation. Portable sensors (HandHeld BUD, MPV, TEMTADS 2x2) were deployed in a treed area not feasible for vehicular sensors and the MetalMapper was deployed in an open field area. The Combined Area was surveyed by all instruments. 20
Figure 2: Overview of workflow applied to each dataset from Camp Beale. 21
Figure 3: Sensors deployed in the Portable area at Camp Beale included, from left to right, Geonics EM61, TEMTADS 2x2x3, HandHeld BUD, MPV. 21
Figure 4: Feature space plot for the EM61 sensor in the Camp Beale Portable Area. 22
Figure 5: ROC curves for the EM61 sensor in the Camp Beale Portable Area. Results are shown both for fuzes considered clutter (left) and fuzes considered UXO (right). 22
Figure 6: Feature space plot for the HandHeld BUD sensor in the Camp Beale Portable Area. . 23
Figure 7: ROC curves for the HandHeld BUD. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right). 24
Figure 8: Feature space plot for the TEMTADS 2x2x3 sensor in the Camp Beale Portable Area. 25
Figure 9: ROC curves for the TEMTADS 2x2x3 sensor. Results are shown both for fuzes considered clutter (left) and fuzes considered UXO (right). 26
Figure 10: Feature space plot for the MPV sensor in the Camp Beale Portable Area. 27
Figure 11: ROC curves for the MPV. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right). 28
Figure 12: MetalMapper sensor collecting data at Camp Beale (left). Sensor consists of three orthogonal transmitters and seven multi-static receiver cubes (on right). 30

Figure 13: Feature space for the EM61 sensor in the Camp Beale Open Area. 31
Figure 14: ROC curves for the EM61 sensor in the Camp Beale Open Area. Results are shown both for fuzes considered clutter (left) and fuzes considered UXO (right). 31
Figure 15: Training data for both TOI 1712 and 1509 both exhibited excellent matches to the primary polarizability of a library item but poorly constrained secondary polarizabilities..... 32
Figure 16: Feature space for the CH2M Hill MetalMapper sensor in the Camp Beale Open Area. 33
Figure 17: ROC curves for the CH2M Hill MetalMapper. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right). 34
Figure 18: Feature space for the Parsons MetalMapper sensor in the Camp Beale Open Area. .. 35
Figure 19: ROC curves for the Parsons MetalMapper. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results for the SVM diglist are on the bottom row (fuzes as clutter on left, fuzes as TOI on right). 35
Figure 20: ROC curves for the prescreened CH2M Hill MetalMapper data. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right). 36
Figure 21: ROC curves for the prescreened Parsons MetalMapper data. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right). 37
Figure 22. Project management hierarchy showing Sky Research personnel in blue and UBC-GIF personnel in green. The hierarchy is split between the development and execution components. 39
Figure 23. Dig order for PNN classifier applied to Beale EM-61 open field data. The number in the top right of each subplot indicates the first *N* labeled test items found during digging, displayed as black markers in the test data. Digging in the first stage was terminated after 1200 items (red subplot). Horizontal red line indicates a decay rate cut-off of 0.08..... A-1
Figure 24. Retrospective ROC curves for classification with EM-61 data, Beale handheld area. A-2
Figure 25. Comparison of polarizability size (left) and data amplitude (right) features for EM-61 data, Beale handheld area. A-3
Figure 26. Prescreened feature vectors (red markers) for Beale open area. Target highlighted in magenta is missed seed target 2277. ..... A-3
Figure 27. Groundtruth (left) and Parsons MetalMapper polarizabilities (right) for seed item 2277. Target response is likely dominated by wires and recovered polarizabilities do not match with 37mm items..... A-4
Figure 28. Example of an unrealistic MOI model (anomaly 1588; scrap). The first model of the MOI (model 2) provides (by far) the best fit to the reference polarizabilities (misfit = 0.426), but the predicted depth of 1.05m and high amplitude of the polarizabilities, especially in relation to the relatively weak EM61 anomaly, are indicators that this model is an artifact of the MOI process. Accordingly, this model was failed during QC. B-2

Figure 29. Distribution of all passed models in decay(t1,t29) versus size(t1) feature space, where size(t1) is the total polarizability measured at the first time channel (t1=0.106ms), and decay(t1,t29) is size(t1)/size(t29) where t29=2.006ms. Circles, squares and triangles correspond to SOI and MOI (models 1 and 2), respectively. Labeled stars show location of library reference items in feature space. B-3
Figure 30. Example of use of training data selection tool. A polygon (heavy black line) is drawn in feature space. Clusters of items with self-similar polarizabilities are automatically found. In this case two clusters were found within the polygon; one is visible (solid feature symbols surrounded by broken grey line). Polarizabilities for this cluster are shown in Figure 31. ..... B-4
Figure 31. Polarizabilities for the cluster shown in Figure 30. Training data were requested for Anomaly 138 (check marked); ground truth showed that this anomaly corresponds to a 105mm TOI. B-5
Figure 32. Cluster with polarizabilities different from those of all reference items. Items in the cluster are solid feature symbols surrounded by broken grey line. Polarizabilities for this cluster are shown in Figure 33. ..... B-6
Figure 33. Polarizabilities for the cluster shown in Figure 32. The polarizabilities are not UXO-like and based on past experience were recognized as representative of horseshoes. Training data were not requested for any of these items. B-7
Figure 34. Cluster with polarizabilities similar to those of a sphere. Items in the cluster are solid feature symbols surrounded by broken grey line. Inset shows close up of the polygon surrounding the cluster. Polarizabilities for this cluster are shown in Figure 35. ..... B-8
Figure 35: Polarizabilities for the cluster shown in Figure 34. Training data was requested for the three checked items: anomaly numbers 177, 1648 and 1901. Two of these (177 and 1648) turned out to be scrap. Anomaly 1901 is a fuze part, initially considered by the program office to be a TOI. B-9
Figure 36. Dig list tool graphical user interface. Features are plotted in decay-size space with each feature color-coded according to its location in the dig list (red earliest; black latest). Features colored white are training items. Inset at top left shows polarizabilities of the selected anomaly (feature surrounded by a black square). The dig list order is based on the weights shown on the right. These can be specified manually or optimal weights can be determined by a search procedure. The latter approach was used for the Beale Parsons MetalMapper data. B-11
Figure 37. Example display of polarizabilities plotted in dig list order. Here, polarizabilities for dig numbers 325 through 360 are shown. Number in top right of each panel is dig number. Numbers in lower left are anomaly number (preceded by "T") and misfit (calculated using all three polarizabilities) to best fitting reference polarizabilities (broken light blue lines). Text in lower left corner is name of closest fitting reference item. Each panel is shaded according to misfit, with larger misfits corresponding to darker shading. Yellow-highlighted dig number (342) is the stop dig point that was chosen for the dig list in which pre-screening was not employed… B-12
Figure 38. Partial ROC curve for the dig list with no pre-screening. Fuzes are considered TOI. B-13
Figure 39. By fitting a bi-normal distribution to the observed partial ROC curve the number of digs necessary to achieve a specified confidence level (that all TOI have been dug) can be estimated. At the 99% confidence level, our partial results suggested that no further digging was necessary. B-13

Figure 40. Decay versus size feature plot for Parsons MetalMapper data. Large symbols show locations of TOI in feature space. Two ISOs that were missed (anomalies 1965 and 2532) are identified. Eleven small fuzes marked with "X" were missed (for the case of fuzes treated as TOI)..... B-14
Figure 41. Final ROC curve for the case of fuzes treated as TOI..... B-15
Figure 42. Polarizabilities of the eleven fuzes that were missed for the case of fuzes considered as TOI..... B-15
Figure 43. Ground truth photos of the eleven fuzes that were missed for the case of fuzes considered as TOI. B-16
Figure 44. Final ROC curve for the case of fuzes treated as clutter. B-17
Figure 45. Polarizabilities and ground truth photos of the two ISOs that were missed. Anomaly 1965 appeared at dig 447; anomaly 2532 appeared at dig 728..... B-17
Figure 46. Retrospective dig list (fuzes considered as clutter) constructed following the same procedure used to construct to the dig list shown in Figure 44, but changing to matching on the primary polarizability after 211 digs. All TOI are dug after 262 digs (126 non-TOI digs)..... B-18
Figure 47. Retrospective dig list (fuzes considered as clutter) based strictly on match to the primary polarizability. All TOI are dug after 260 digs (124 non-TOI digs)..... B-18
Figure 48: Inversion results for target 2529, a 37mm projectile. An excellent fit with high SNR, recovered positions corresponding to EM61 anomaly location..... B-20
Figure 49: Inversion result for target 2136, a 60mm mortar. Results are shown for the single-object inversion. Recovered locations are near the edge of the array at one of the inversion bounds. Polarizabilities are not an excellent match to any of the reference items and one of the secondary polarizabilities is not well constrained. This model was failed as a result (the model from the multi-object inversion corresponding to approximately the same location was also failed for similar reasoning)..... B-20
Figure 50: The QC tool display for target 2136, a 60mm mortar. Results are shown for the multi-object inversion. Recovered locations are in agreement with the EM61 anomaly. Recovered polarizabilities produce an excellent match to the 60mm reference item and this model was passed as a result..... B-21
Figure 51: The Training data tool was used to assist the analyst in identifying target clusters in a size/decay feature space..... B-22
Figure 52: The Training data tool identified this cluster of self similar targets which had 3 approximately equal polarizabilities. Training data requests for these items identified them as fuzes. B-23
Figure 53: Training data requests for TOI where primary polarizability was an excellent match to reference item however secondary polarizabilities were not well recovered. B-24
Figure 54: Groundtruth from the Stage 2 diglist revealed a small 5cm fuze that was not identified earlier through clustering of similar targets in feature space. The L1 polarizability from the multi object inversion was used to search for similar items once the target had been identified. B-26
Figure 55: ROC curves for CH2M Hill MetalMapper data for the case where fuzes were considered to be TOI (top) and with fuzes designated clutter (bottom). B-27
Figure 56: Polarizabilities for the three missed items when fuzes were considered TOI. In addition to the ISO (see Figure 57), 5cm and 6cm fuzes were also not identified as TOI B-28
Figure 57: Summary showing key portions of the QC tool for missed ISO (Target 1786). Polarizabilities recovered from both single and multi object inversions do not produce a good match to reference ISO polarizabilities. Target location indicates the recovered location was near

the corner of the MetalMapper array at the inversion bounds and should have been flagged as can't analyze. B-29
Figure 58. Size-decay space for Beale MetalMapper Parsons data. Assumed non-TOI are test feature vectors with maximum total polarizability misfit with training TOI. C-2
Figure 59. Decision surface for two-stage SVM classifier applied to Parsons MetalMapper data. An SVM prediction of 1 indicates a high likelihood of TOI. C-2
Figure 60. SVM decision statistic f_{SVM} for stage 1 (all polarizabilities) SVM classifier applied to Beale Parsons MetalMapper tests data. Marker indicates point in dig list at which we switch to stage 2 (total polarizability) classifier. C-3
Figure 61. Predicted binormal ROC and 99% confidence interval for first stage groundtruth from Beale Parsons MetalMapper data. An additional 19 digs (from 546 to 565) are required to test the null hypothesis. C-5
Figure 62 (a) Effect of retraining on dig order for full test data set (b) Final dig order after re-ranking to minimize undesired increases in rank. No re-ranking is applied for the first 565 digs (indicated by vertical red line) because these items have already been labeled in the first stage of digging. Unlabeled targets bumped to the start of the dig order are shown in red. C-6
Figure 63. Retrospective comparison of SVM classification strategies applied to Camp Beale Parsons MetalMapper data. C-7
Figure 64: Schematic of the HandHeld BUD D-1
Figure 65: Ordnance of interest: (a) 37 mm projectile. (b) 60 mm mortar. (c) 81 mm mortar. (d) 105 mm HEAT. D-3
Figure 66: Recovered polarizations from the test-pit data. (a) 37 mm projectile. (b) 60 mm mortar. (c) 81 mm mortar. (d) 105mm projectile. In each plot, the same item was positioned at various depths and orientations (horizontal, vertical (nose up and down), tilted at 45^0). D-4
Figure 67: Measurement template in Camp Beale. D-4
Figure 68: Scatter plot of size and decay for BUD data sets at Camp Beale. Based on the minimum (mfp) misfit of polarizabilities recovered from the test data to the references., we have the 3 rough partitions of the data points for an initial analysis. There are the two potential clusters (dashed ovals) visually observed and are isolated from the ordnance features of 37mm projectiles, 60mm mortars, 81 mm projectiles, and 105 mm projectiles. D-6
Figure 69: Recovered polarizations for the training data. Top row: BE-680 (ISO), BE-411 (37mm projectile), BE-355 (81mm mortar), BE-903 (81mm mortar). Middle row: BE-905 (fuze), BE-798 (fuze), BE-143(fuze), BE-669 (fuze). Bottom row: BE-301 (munitions debris), BE-458 (horseshoe), BE-378 (frag), BE-367 (frag). D-7
Figure 70: BE-669. The Ground truth and the recovered polarizabilities from a two-object inversion (2OI). D-7
Figure 71: BE-538. The recovered polarizabilities, the ground truth picture, and the fits between the observed (green) and predicted (blue) data. D-9
Figure 72: BE-854. The Ground truth and the recovered polarizabilities that are closet to those of ISO was dug at stage 3. ..... D-10
Figure 73: ROC curves at stage 1 and 3 for BUD at Camp Beale. D-11
Figure 74: Some examples of the dug anomalies that can cause classification ambiguity. The recovered polarizabilities and the ground truth (GT). The gray curves represent the closest reference polarizations in each plot. D-12
Figure 75: The recovered polarizabilities and the ground truth of the missed small fuze targets. The gray curves represent the reference polarization of fuze in the library. D-13

Figure 76: Figure 13. Feature plot of size and decay for BUD data sets at Camp Beale for H-TOI (high TOI) set and L-TOI (likely or low TOI) set at stage-1 dig list. The green crosses denote anomalies in L-TOI with their size feature smaller than 1. ..... D-14
Figure 77: Depth-range plot for 37mm projectiles, 60mm mortars, 81mm mortars, 105mm projectiles, fuzes, and ISO identified in the H-TOI (high TOI) set and the crosses represent the identified items in the L-TOI (likely or low TOI) set at stage-1 dig list. The green crosses denote anomalies in L-TOI with their size feature smaller than 1. ..... D-15
Figure 78: Recovered locations for the Parsons MetalMapper Data set. The scatter plot of location errors in (a) suggest two clusters. In (b) the geographic distribution of anomalies are plotted. The anomalies within the red dashed line in (a) are plotted in (b) with a blue circle....E-2
Figure 79: Recovered locations from Parsons ((a)-(b) and CH2M HILL MetalMapper data ((c)-(d)). All Open Area anomalies (including those in the shared area) are plotted. E-3
Figure 80: Positional errors for the Parsons MetalMapper data. The top row ((a)-(b)) contain anomalies in the western region of the Open Area. The bottom row ((b)-(c)) contain anomalies in the eastern region of the Open Area, which shares anomalies with the Portable Area. E-4
Figure 81: Positional errors for the CH2M Hill MetalMapper data. The top row ((a)-(b)) contain anomalies in the western region of the Open Area. The bottom row ((b)-(c)) contain anomalies in the eastern region of the Open Area, which shares anomalies with the Portable Area. E-5
Figure 82: Depth errors for the Parsons MetalMapper data (top row) and CH2MHill data (bottom row). E-6
Figure 83: Depth analysis for portable sensors…… …… ……E-7

Acronyms

BUD	Berkeley UXO Discriminator
EM	Electromagnetic
EMI	Electromagnetic induction
ESTCP	Environmental Security Technology Certification Program
FAR	False alarm rate
GPS	Global Positioning Systems
HEAT	High explosive anti-tank
IDA	Institute for Defense Analyses
IMU	Inertial Measurement Unit
ISO	Industry standard object
IVS	Instrument verification strip
MEC	Munitions and explosives of concern
MPV	Man Portable Vector
MOI	Multi object inversion
MTADS	Multi-Sensor Towed Array Detection System
PI	Principal Investigator
PNN	Probabilistic Neural Network
POC	Point of Contact
QC	Quality Control
ROC	Receiver operating characteristic
RTS	Robotic Total Station
SERDP	Strategic Environmental Research and Development Program
SLO	San Luis Obispo
SNR	Signal to noise ratio
SOI	Single object inversion
SVM	Support vector machine
TEM	Time-domain electromagnetic
TEMTADS	Time Domain Electromagnetic Towed Array Detection System
TOI	Target of interest
UBC-GIF	University of British Columbia – Geophysical Inversion Facility
UBC	University of British Columbia
USACE ERDC	United States Army Corps of Engineers- Engineering Research and Development Center
UXO	Unexploded ordnance

1.0 INTRODUCTION

1.1 BACKGROUND

Significant progress has been made in discrimination technology. To date, testing of these approaches has been primarily limited to test sites with only limited application at live sites. Acceptance of discrimination technologies requires demonstration of system capabilities at real UXO sites under real world conditions. Any attempt to declare detected anomalies to be harmless and requiring no further investigation will require demonstration to regulators of not only individual technologies, but an entire decision making process.

To demonstrate the viability of advanced detection and discrimination technologies, ESTCP has now conducted four UXO classification studies. The results of the first demonstration, at the former Camp Sibert, Alabama were very encouraging. Although conditions were favorable at this site, including a single target-of-interest (4.2-in mortar) and benign topography and geology, all of the demonstrated classification approaches were able to correctly identify a sizable fraction of the anomalies as arising from non-hazardous items that could be safely left in the ground. Of particular note, the contractor EM-61-MK2 cart survey with analysis using commercially available methods correctly identified more than half the targets as non-hazardous.

To build upon the success of this first study, ESTCP expanded the program to include a second study at a site with more challenging topography and a wider mix of targets-of interest. A range at the former Camp San Luis Obispo (SLO), California, was selected for this demonstration. We again found that, with appropriate use of classification metrics applied to production quality EM-61 data, it was possible to significantly reduce the number of clutter items excavated without missing any targets of interest (TOI). Furthermore, the next generation of EM sensors, when deployed in a cued-interrogation mode, produced significant additional reductions in the number of clutter items excavated. These sensors could also usually distinguish between different UXO types. A third ESTCP demonstration study was conducted in 2010 at Camp Butner, North Carolina. The site had very little topographic relief but required classification between small targets of interest (37mm projectiles and M48 fuzes) and metallic debris of similar size. Targets were also distributed with a higher density than previously encountered.

The latest ESTCP demonstration study was conducted in 2011 at Camp Beale, California. The site had very little topographic relief but required classification between small targets of interest (37mm projectiles and M48 fuzes) and metallic debris of similar size. Targets were also distributed with a higher density than previously encountered. This demonstration report describes the data processing, feature extraction and classification that were conducted by Sky Research (SKY) and the University of British Columbia (UBC) on the Camp Butner data sets.

1.2 OBJECTIVE OF THE DEMONSTRATION

The objectives of this demonstration were to perform data modeling, classification, and classification using electromagnetic (EM) data collected by the various data collection demonstrators participating in the study. Specifically, we processed the following datasets collected at Camp Beale:

1) EM-61 cart data;

2) Man-Portable Vector (MPV) EMI sensor cued interrogation data;
3) TEMTADS 2x2x3 cued interrogation array data;
4) Data collected using the Handheld Berkeley UXO Discriminator (Handheld BUD) deployed in a cued interrogation mode;
5) MetalMapper cued interrogation data acquired by Parsons; and
6) MetalMapper cued interrogation data acquired by CH2M Hill;

Specific processing tasks were as follows:

1) Feature Extraction: inversion of all electromagnetic data sets with a three dipole instantaneous amplitude model
2) Classification: production of the following ranked dig sheets for the Beale Portable Area :

    a) EM-61 statistical: Discrimination achieved using features derived from the data amplitude versus decay feature space;
    b) HandHeld Berkley UXO Discriminator (BUD) polarization match: A digsheet was produced based on how well the recovered polarization matched the polarizations in a library of ordnance items expected at the site;
    c) HandHeld (BUD) Support Vector Machine (SVM): We provided an alternative ranking of the HandHeld BUD based on SVM;
    d) Man Portable Vector (MPV) polarization match: As in b) but for the MPV; and
    e) Man Portable Vector (MPV) SVM: As in c) but for the MPV.

As well, the following ranked dig sheets were produced for the Beale Open Area:

    f) CH2MHill MetalMapper polarization match: A digsheet was produced based on how well the recovered polarization matched the polarizations in a library of ordnance items expected at the site;
    g) CH2MHill MetalMapper SVM: We provided an alternative ranking of the CH2MHill MetalMapper based on SVM;
    h) Prescreened CH2MHill MetalMapper polarization match: As in f) except 231 high confidence non-TOI items identified in EM-61 data were assumed non-TOI and not analyzed;
    i) Prescreened CH2MHill MetalMapper SVM: As in g) except 231 high confidence non-TOI items identified in EM-61 data were assumed non-TOI and not analyzed;
    j) Parsons MetalMapper polarization match: As in f) but for Parsons MetalMapper;
    k) Parsons MetalMapper SVM: As in g) but for Parsons MetalMapper;

l) Prescreened Parsons MetalMapper polarization match: As in h) but for Parsons MetalMapper; and

m) Prescreened Parsons MetalMapper SVM: As in i) but for Parsons MetalMapper.

Thus we produced a total of thirteen ranked dig sheets using a variety of different methods and sensor types.

1.3 REGULATORY DRIVERS

Refer to the Program Office demonstration plan for a discussion of regulatory drivers.

2.0 TECHNOLOGY

2.1 TECHNOLOGY DESCRIPTION

Magnetic and EM methods represent the main sensor types used for detection of UXO. Over the past 10 years, significant research effort has been focused on developing methods to discriminate between hazardous UXO and non-hazardous scrap metal, shrapnel and geology (e.g. Bell et al., 2001; Pasion et al., 2007; Tantum et al., 2008; Liao and Carin, 2009). The most promising classification methods typically proceed by first recovering a set of parameters that specify a physics-based model of the object being interrogated. For example, in time-domain electromagnetic (TEM) data, the parameters comprise the object location and the polarization tensor (typically two or three collocated orthogonal dipoles along with their orientation and some parameterization of the time-decay curve). For magnetics, the physics based model is generally a static magnetic dipole. Once the parameters are recovered by inversion, a subset of the parameters is used as feature vectors to guide either a statistical or rule-based classifier.

Magnetic and EM phenomenologies have different strengths and weaknesses. Magnetic data are simpler to collect, are mostly immune to sensor orientation and are better able to detect deeper targets. EM data are sensitive to non-ferrous metals, are better at detecting smaller items and are able to be used in areas with magnetic geology. Only EM data was acquired at Camp Beale and in the remainder of this report we therefore focus on EM sensing and processing..

There are three key elements that impact the success of the UXO classification process described in the previous paragraphs:

1) Creation of a map of the geophysical sensor data: This includes all actions required to form an estimate of the geophysical quantity in question (i.e. amplitude of EMI response at a given time-channel) at each of the visited locations. The estimated quantity is dependent on the following:
    a. Hardware, including the sensor type, deployment platform, position and orientation system and the data acquisition system used to record and time-stamp the different sensors;
    b. Survey parameters such as line spacing, sampling rate, calibration procedures etc.;
    c. Data processing such as merging of position/orientation information with sensor data, noise and background filtering applied;
    d. The background environment including geology, vegetation, topography, cultural features, etc.; and
    e. Depth and distribution of ordnance and clutter.
2) Anomaly selection and feature extraction: This includes the detection of anomalous regions and the subsequent extraction of a polarization tensor model for each anomaly.
3) Classification of anomalies: The final objective of the demonstration is the production of a dig sheet with a ranked list of anomalies. This will be achieved via statistical classification which will require training data to determine the attributes of the UXO and non-UXO classes.

The focus of this demonstration is on the further testing and validation of the methodologies for 2) and 3) above that have been developed in UXOLab jointly by Sky Research and the University of British

Columbia-Geophysical Inversion Facility (UBC-GIF). We now describe each of the three key elements of the technology as identified above.

2.1.1 Creation of a Map of Geophysical Sensor Data

At Camp Beale, each of the demonstrators provided filtered, geo-located geophysical data. No additional pre-processing was applied to the data sets.

2.1.2 Anomaly Selection and Feature Extraction

At this point in the process flow, there is a map of each of the geophysical quantities measured during the survey. The next step in the process is detection of anomalous regions followed by the extraction of features for each of the detected items. For this demonstration, targets have been picked from the EM-61 cart data by the demonstrator, no additional picks were made by Sky/UBC.

In the EMI method, a time varying field illuminates a buried, conductive target. Currents induced in the target then produce a secondary field that is measured at the surface. EM data inversion involves using the secondary field generated by the target for recovery of the position, orientation, and parameters related to the target's material properties and shape. In the UXO community, the inverse problem is simplified by assuming that the secondary field can be accurately approximated as a dipole. In general, TEM sensors use a step off field to illuminate a buried target. The currents induced in the buried target decay with time, generating a decaying secondary field that is measured at the surface. The time-varying secondary magnetic field $\mathbf{B}(t)$ at a location $\mathbf{r}$ from the dipole $\mathbf{m}(t)$ is computed as:

$$\mathbf{B}(t) = \frac{\mu_o}{4\pi r^3}\mathbf{m}(t)\cdot\left(3\hat{\mathbf{r}}\hat{\mathbf{r}} - \overline{\overline{\mathbf{I}}}\right) \tag{1}$$

where $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$ is the unit-vector pointing from the dipole to the observation point, $\mathbf{I}$ is the 3 x 3 identity matrix, $\mu_o = 4\pi \times 10^{-7}$ H/m is the permittivity of free space and $r = |\mathbf{r}|$ is the distance between the center of the object and the observation point.

The dipole induced by the interaction of the primary field $\mathbf{B}_o$ and the buried target is given by:

$$\mathbf{m}(t) = \frac{1}{\mu_o}\overline{\overline{\mathbf{M}}}(t)\cdot\mathbf{B}_o \tag{2}$$

The induced dipole is the projection of the primary field $\mathbf{B}_o$ onto the target's polarizability tensor **M(t)**. The polarizability tensor is assumed to be symmetric and positive definite and so can be decomposed as

$$\boldsymbol{M}(t) = \boldsymbol{A}^T\mathbf{L}(t)\boldsymbol{A} \tag{3}$$

with $\boldsymbol{A}$ an orthogonal matrix which rotates the coordinate system from geographic coordinates to a local, body centered coordinate system. The diagonal eigenvalue matrix $\mathbf{L}(t)$ contains the principal polarizabilities $L_i(t)$ (i = 1, 2, 3), which are assumed to be independent of target orientation and location. Features derived from the dipole model, in particular amplitude and decay of the principal polarizabilities, have been successfully used to discriminate between targets of interest and non-hazardous metallic clutter. These parameters are useful because, to first order, a conductor can be modeled as a simple LR loop which

is inductively coupled to transmitters and receivers on the surface. The current response of this loop is a decaying exponential which is fully described by an amplitude and time constant (West and Macnae, 1991). The TEM dipole model generalizes this simple circuit model to account for target size and shape. This latter property is represented by the principal polarizabilities, which decay independently in time and are approximately aligned with the semi-major and minor axes of the target.

Equal transverse (secondary and tertiary) polarizabilities indicate an axisymmetric target. Most ordnance can be treated as bodies of revolution (Shubitidze et al., 2002), and so equality of transverse polarizabilities has been proposed as a useful feature for discriminating between TOI and irregularly-shaped clutter. However, in practice it has been difficult to reliably estimate target shape using data from mono-static, vertical-component sensors conventionally deployed for UXO detection. This is because monostatic data often cannot adequately interrogate the transverse response of buried targets.

Recent advances in TEM sensor technology for UXO detection have helped address these limitations. For example, the MetalMapper sensor is comprised of an array of 7 receivers that measure 3 orthogonal components of the secondary field generated by 3 orthogonal transmitter loops that are fired sequentially. This multi-static, and multi-transmitter configuration provides a very rich data set which is better able to constrain target depth and transverse polarizabilities than a mono-static sensor.

When solving parametric inverse problems, it is usually sufficient to minimize a data norm quantifying the misfit between observed (dobs) and predicted data

$$\phi_d = \left\| V_d^{-1/2} \left(d^{obs} - d^{pred} \right) \right\|^2 \qquad (4)$$

with $d^{pred} = F(m)$ generally a nonlinear functional of the model m, and $Vd^{-1/2}$ a (usually diagonal) covariance matrix specifying estimated errors on the data. Bound constraints are also typically imposed to ensure that physically reasonable model parameters are obtained (e.g. polarizabilities should be positive). In the case of TEM data, the model is parameterized in terms of target location and orientation, as well as principal polarizabilities at each time channel. Equation 4 is minimized by first estimating the target location, followed by estimating of polarizabilities at each time channel. Decoupling the time channels in this way makes the inversion less sensitive to the specified uncertainties, but produces polarizabilities that are less smooth as a function of time.

2.1.3 Classification of Anomalies

At this stage in the process, we have feature vectors for each anomaly and need to decide which items should be excavated as potential UXO. For this demonstration, we employ two methods for classifying anomalies: (1) Library based techniques and (2) Statistical classification methods. Library based methods compare the recovered polarizabilities from a target to polarizabilities for a library of targets that may be found at the site. Anomalies are ranked according to the minimum misfit to the library items. Statistical classification techniques have proven to be very effective at classification at various test sites (e.g. Billings, 2010). Statistical classifiers have been applied to a wide variety of pattern recognition problems, including optical character recognition, bioinformatics and UXO classification. Within this field there is an important dichotomy between supervised and unsupervised classification. Supervised classification makes classification decisions for a test set comprised of unlabeled feature vectors. The classifier performance is optimized using a training data set for which labels are known. In unsupervised classification there is only a test set; labels are unknown for all feature vectors. Most applications of

statistical classification algorithms to UXO discrimination have used supervised classification; the training data set is generated as targets are excavated. More recently, unsupervised methods have been used to generate a training data set that is an informative sample of the test data (Zhang, 2004). In addition, semi-supervised classifiers, which exploit both labeled data and the topology of unlabelled data, have been applied to UXO classification in one study (Zhang, 2004).

A popular and proven algorithm employed in the machine learning community for supervised classification and regression problems is the support vector machine (SVM). For the Camp Butner data sets we used this classifier to produce a number of ranked dig sheets, and so here we provide a brief development of the algorithm. We begin with a training set of *M* feature vectors in an *N*-dimensional feature space

$$\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_M\}, \quad x_i \in \boldsymbol{R}^N. \tag{5}$$

In a simple two-class case, the feature vectors have associated labels

$$\boldsymbol{Y} = \{y_1, y_2, \ldots y_M\}, \quad y_i \in \{-1,1\}. \tag{6}$$

Define a support plane for a given class to be a line (or hyperplane in higher dimensions) so that all feature vectors in that class fall to one side of the support plane, as illustrated in Figure 0.1. The margin between classes is the perpendicular distance between support planes. The SVM algorithm then tries to find an optimal decision plane by maximizing the margin, subject to the constraint that the data are classified correctly. Referring to Figure 0.1, the decision function is

$$f(\boldsymbol{x}) = \boldsymbol{w} \cdot \boldsymbol{x} + b \tag{7}$$

with $\boldsymbol{w}$ the normal vector, and *b* an offset. The decision boundary is defined by points for which $f(\boldsymbol{x}) = 0$, and a ranking of feature vectors can be obtained by thresholding on the decision function.

To maximize the margin we then find that we can minimize $\|\mathbf{w}\|^2/2$. It can be shown that this optimization problem can be solved by maximizing the dual problem

$$\mathbf{L}_d = \left[\sum_{i=1}^{M} \alpha_i - \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \alpha_i y_i \left(\mathbf{x}_i^T \mathbf{x}_j\right) y_j \alpha_j\right] \tag{8}$$

With the constraints

$$\sum_{i=1}^{M} \alpha_i y_i = 0. \tag{9}$$

The dual problem gives us a solution for the normal to the hyperplane

$$\boldsymbol{w} = \sum_{i=1}^{M} \alpha_i y_i \mathbf{x}_i. \tag{10}$$

The solution has the property that the majority of coefficients α_i are zero, with nonzero values corresponding to feature vectors lying on the support planes (these are termed the support vectors).

A more general formulation of the SVM allows for nonlinear decision boundaries with overlapping classes. The idea is to map the feature data to a higher-dimensional space where the data become separable and to construct the optimal separating hyperplane in this space. In equation 8 we need only evaluate the inner product of the feature vectors. Hence to construct the decision boundary in another space we need only compute inner products in that space. The optimization problem becomes

$$\mathbf{L}_d = \left[\sum_{i=1}^{M} \alpha_i - \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \alpha_i y_i K\left(\mathbf{x}_i, \mathbf{x}_j\right) y_j \alpha_j\right] \quad (11)$$

with the same constraint as before (equation 9). The most common choices of kernel functions $K(\mathbf{x}_i, \mathbf{x}_j)$ for nonlinear SVMs are radial basis functions and polynomials. While the choice of kernel functions is often arbitrary, practical experience has shown that SVMs trained with different kernels often have a large percentage of support vectors in common.

Kernel function	$K(\mathbf{x}_i, \mathbf{x}_j)$
Radial Basis function	$\exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$
Degree p polynomial	$(x_i \cdot x_j + 1)^p$

Table 0.1: Kernel functions commonly used with nonlinear support vector machines.

The complexity of the decision boundary is controlled by the kernel parameter (σ in the case of the radial basis function). A small kernel parameter results in a decision boundary that fits the training data closely, but may not generalize well to the test data.

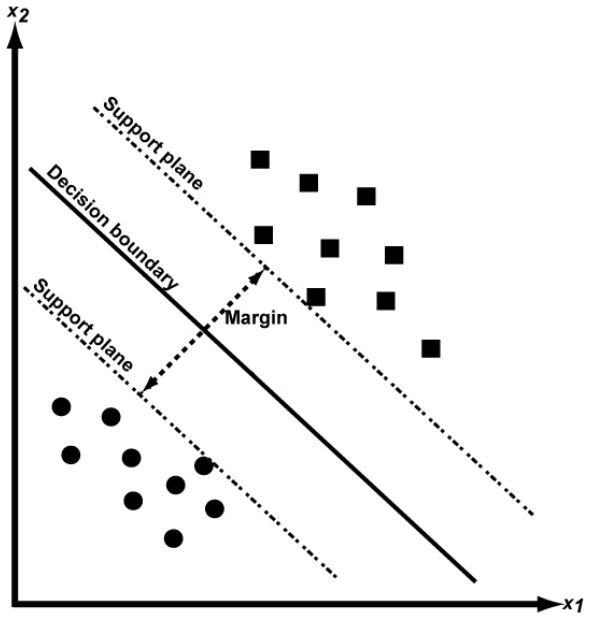


Figure 0.1: SVM formulation for constructing a decision boundary. The decision boundary bisects support planes bounding the classes.

2.2 UXOLab Software

The methodologies for data processing, feature extraction, and statistical classification described above have been implemented within the UXOLab software environment, which was used for this demonstration. UXOLab is a Matlab based software package developed over a six year period at the

UBC-GIF, principally through funding by the United States Army Corps of Engineers-Engineering Research and Development Center (USACE ERDC) (DAAD19-00-1-0120). Over the past five years, Sky Research and UBC-GIF have considerably expanded the capabilities of the software.

2.3 PREVIOUS TESTING OF THE TECHNOLOGY

Table 1 provides a list of some of the previous tests conducted of the underlying data processing and interpretation methodology that will be used in this demonstration.

Table 1. Previous Inversion/Classification Testing

Inversion/Classification Test	Description	Results
Demonstration Site: Camp Sibert		
Geonics EM-61 cart, MTADS EM-61 array, MTADS mag array, and EM-63 single and cooperative inversions. EM-63 cued interrogations were positioned by a Leica TPS 1206 RTS with orientation information provided by a Crossbow AHRS 400 IMU. The objective of the surveys was the classification of a large target (4.2-in mortars). The site was unusual in that the primary munition known to have been used was the 4.2-in mortar, thus providing a site where the classification is a case of identifying a single large target amongst smaller pieces of mortar debris and clutter.	For the EM-61, 3-dipole instantaneous amplitude models were fit to the available 3 time-channels, while for the EM-63, 3-dipole Pasion-Oldenburg models were recovered from the 26 time-channel data. MTADS and EM-63 data were also cooperatively inverted. Parameters of the dipole model were used to guide a statistical classification.	The results for all sensor combinations were excellent, with just one false negative for the EM-63 when inverted without cooperative constraints. When inverted cooperatively, the EM-63 cued interrogation was the most effective discriminator. All 33 UXO were recovered with 25 false alarms (16 of these were in the "can't analyze" category). Not counting the "can't analyze" category, the first 33 recommended excavations were all UXO. The MTADS and MTADS cooperatively inverted were also very effective at discrimination, with all UXO recovered very early in the dig list. The MTADS data set suffered from a high number of false alarms due to anomalies with a geological origin. In addition, the operating point was very conservative and many non-UXO were excavated after recovery of the last UXO in the dig list. The results from the EM-61 cart were also very good, although 24 false-positives were required to excavate all 105 UXO. The lower data quality of the EM-61 cart resulted in a larger number of \can't analyze" anomalies over metallic sources than the MTADS.
Demonstration Site: San Luis Obispo		
Detection mode surveys were: MTADS magnetometer and EM61 arrays, Geonics EM61 cart and Man-Portable Simultaneous EMI and Magnetometer System (MSEMS) cart. Cued interrogation mode TEMTADS, MetalMapper and Berkeley UXO Discriminator (BUD) surveys were also conducted. At this site the objective was to identify TOI from a number of different target classes: primarily 60 mm, 81 mm, 4.2" mortars, 2.36" rockets, and one each of 37 mm, 3" and 5" projectiles. The site had significant topographic relief.	For all TEM data sets, 3-dipole instantaneous amplitude models were fit to the available time-channels. Single dipoles were fit to targets in magnetics data sets, but cooperative inversion was not used at this site. For detection data sets a threshold on the rate of decay of primary polarizabilities was used to rank targets. Dig sheets for cued interrogation data sets were generated using statistical classifiers trained on size/decay features, as well as with library methods and an ``expert" method based on the judgment of an analyst.	Magnetometer detection and classification performance at this site was quite poor. The EM-61 production datasets were much more effective than magnetics. Size estimated from the recovered polarizations was not an effective classification metric due to the small-size of the 60 mm mortars and the inability to accurately constrain depth. However, the time-decay rate estimated from the recovered polarizabilities provided an effective ranking scheme. The EM61 cart performance was marginally better than the MSEMS cart and MTADS EM61 array. For the TEMTADS data the library method was the most effective with 204 of 206 TOI recovered along with 131 of 1076 non-TOI. The two false negatives were the rocket motor pieces declared non-TOI by all cued-interrogation methods and a 60 mm mortar with a target response that overlapped with some nearby clutter. The other classification methods were also effective, generating between 2 to 4 false-negatives.

Inversion/Classification Test	Description	Results
		The library method was again most effective for the MetalMapper data with the excavation of 203 of 204 TOI and 175 of 1205 non-TOI. Correct classification of ordnance type was also achieved with up to 99% accuracy achieved with statistical classification.
Demonstration Site: Camp Butner		
Detection surveys with the EM-61 cart and MetalMapper dynamic sensors were conducted. Detected targets from the EM-61 data were revisited with MetalMapper static and TEMTADS cued interrogations. TOI at the site included 105mm HEAT, 37 mm, and M48 fuses. Topographic relief was benign, but there was a significant amount of clutter similar in size and shape to 37 mm.	For all TEM data sets, 3-dipole instantaneous amplitude models were fit to the available time-channels. For detection data sets a threshold on the rate of decay of primary polarizabilities was used to rank targets. Dig sheets for cued interrogation data sets were generated using statistical classifiers trained on size/decay features, as well as with library methods and an ``expert'' method based on the judgment of an analyst.	Thresholding on time-decay rate of estimated polarizabilities estimated from EM-61 data performed quite poorly, with a 0.92 FAR. This is likely attributable to poor depth estimation for small targets. The MetalMapper dynamic data produced reliable depth estimates, but did not measure sufficiently late in time to provide separation between TOI and non-TOI polarizabilities. FAR was approximately 0.7 for this sensor Excellent classification performance was achieved with the TEMTADS data: all 171 TOI were found with a FAR of only 5%. MetalMapper static performance was similar to TEMTADS, but produced a much higher FAR (78%) owing to outlying TOI attributable to faulty data.

2.4 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY

The main advantage of the technology is a potential reduction in the number of non-hazardous items that need to be excavated, thus reducing the costs of UXO remediation. Advantages of UXOLab and the algorithms within the package include:

- All the functionality required to process raw geophysical data, detect anomalous regions, and perform geophysical inversion and classification.
- Algorithms for inverting magnetic and TEM data sets both separately and cooperatively using a number of different polarization tensor formulations.
- Extensive set of algorithms for rule-based and statistical classification algorithms.
- Configuration in a modular fashion, so that as new sensor technologies become available (e.g. new TEM systems with multi-component receivers etc), the inversion functionality will be immediately available to those new sensor systems.

The principal disadvantage is that UXOLab is written in Matlab and has not been configured for general use by contractors and non-specialists. However, as part of ESTCP 1004 we are presently working on transitioning our inversion algorithms to an API that will be generally accessible.

3.0 PERFORMANCE OBJECTIVES

The performance objectives for this demonstration are summarized in Tables 2 and 3. There are objectives for both the data collection and data analysis demonstrators.

The first three analysis objectives refer to the classification part of the demonstration with the first two referring to the best results from each approach in a retrospective analysis and the third addressing how well each demonstrator is able to specify the correct threshold in advance. The final two objectives refer to the feature extraction part of the demonstration.

3.1 OBJECTIVE: MAXIMIZE CORRECT CLASSIFICATION OF MUNITIONS

This is one of the two primary measures of the effectiveness of the classification approach. By collecting high quality data and analyzing those data with advanced parameter estimation and classification algorithms we expect to be able to classify the targets with high efficiency. This objective concerns the component of the classification problem that involves correct classification of items-of-interest.

3.1.1 Metric

The metric for this objective is the number of items on the master anomaly list that can be correctly classified as munitions by each classification approach.

3.1.2 Data Requirements

Each demonstrator will prepare a prioritized dig list for the targets on the master anomaly list. IDA personnel will use their scoring algorithms to assess the results.

3.1.3 Success Criteria

The objective will be considered to be met if all of the items of interest are correctly labeled as munitions on the prioritized anomaly list.

Table 2. Performance Objectives for the Camp Beale Portable Area (fuzes as clutter)

Performance criterion				EM-61	HandHeld BUD		MPV		TEMTADS 2x2x3
Performance Objective	**Metric**	**Data Required**	**Success Criteria**	**Data Amp. Decay**	**SVM**	**Library**	**SVM**	**Library**	**SVM \ Expert**
Maximize correct classification of munitions	Number of targets-of-interest retained	• Prioritized anomaly lists • Scoring reports from (IDA)	Approach correctly classifies all TOI	Yes	Yes	Yes	Yes	Yes	Yes
Maximize correct classification of non-munitions	Number of false alarms eliminated	• Prioritized anomaly lists • Scoring reports from IDA	Reduction of non-munitions > 30% while retaining all TOI (% clutter dug at operating point)	Yes 41%	Yes 22%	Yes 13%	Yes 25%	Yes 37%	Yes 15%
Specification of no-dig threshold	Pd of correct classification and #FA at operating point	• Demonstrator -specified threshold • Scoring reports from IDA	Threshold achieves criteria above (% clutter dug at stop dig point)	100% 805 (100%)	100% 305 (38%)	100% 211 (26%)	100% 323 (40%)	100% 305 (38%)	100% 157 (20%)
Minimize number of anomalies that cannot be analyzed	Number of anomalies that must be classified as "Unable to Analyze"	• Demonstrator target parameters	Reliable target parameters can be estimated for > 90% of anomalies	Yes 100%	Yes 100%	Yes 100%	Yes 100%	Yes 100%	Yes 100%
Correct estimation of target parameters (positions)	Accuracy of estimated target parameters	• Demonstrator target parameters • Results of intrusive investigation	X, Y < 15 cm (1σ) Z < 10 cm (1σ)	NA (x,y) NA (z)	NA (x,y) Yes 2.7cm	NA (x,y) Yes 2.7cm	NA (x,y) Yes 8.8 cm	NA (x,y) Yes 8.8cm	NA (x,y) Yes 7.2cm

Note: Positional error for not calculated for Portable instruments.

Table 3. Performance Objectives for the Camp Beale Open Area (fuzes as clutter)

Performance criterion				EM-61	CH2M Hill				Parsons			
Performance Objective	**Metric**	**Data Required**	**Success Criteria**	**Data Amp. Decay**	**SVM**	**Library**	**Prescreen SVM**	**Prescreen Library**	**SVM**	**Library**	**Prescreen SVM**	**Prescreen Library**
Maximize correct classification of munitions	Number of targets-of-interest retained	• Prioritized anomaly lists • Scoring reports from (IDA)	Approach correctly classifies all TOI	Yes	No 2 FN	No 1 FN	Yes	Yes	Yes	No 2 FN	Yes	No 2 FN
Maximize correct classification of non-munitions	Number of false alarms eliminated	• Prioritized anomaly lists • Scoring reports from IDA	Reduction of non-munitions > 30% while retaining all TOI (% clutter dug at operating point)	Yes 43%	Yes 41%	Yes 60%	Yes 54%	Yes 74%	Yes 80%	Yes 54%	Yes 80%	Yes 45%
Specification of no-dig threshold	Pd of correct classification and #FA at operating point	• Demonstrator - specified threshold • Scoring reports from IDA	Threshold achieves criteria above	100% 1163 (93%)	98.5% 509 (41%)	99.2% 358 (29%)	100% 602 (48%)	100% 315 (25%)	100% 484 (39%)	98.5% 213 (17%)	100% 719 (57%)	100% 176 (14%)
Minimize number of anomalies that cannot be analyzed	No. of anomalies classified as "Unable to Analyze"	• Demonstrator target parameters	Reliable target parameters can be estimated for > 90% of anomalies	Yes 100%	Yes 100%	Yes 100%	Yes 100%	Yes 100%	Yes 100%	Yes 100%	Yes 100%	Yes 100%
Correct estimation of target parameters (positions)	Accuracy of estimated target parameters	• Demonstrator target parameters • Results of intrusive investigation	X, Y < 15 cm (1σ) Z < 10 cm (1σ)	NA (x,y) NA (z)	No A:(16.9,13.7)cm. B:(17.0,12.6)cm Yes 6.8cm				No A:(16.6,13.6)cm. B:(19.4,13.0)cm Yes 7.4cm			

Note: MetalMapper positional errors are split into two parts: "A" is the Open Area without the portion shared with portable instruments, and "B" is the shared area only. See Appendix E for details.

3.2 OBJECTIVE: MAXIMIZE CORRECT CLASSIFICATION OF NON-MUNITIONS

This is the second of the two primary measures of the effectiveness of the classification approach. By collecting high-quality data and analyzing those data with advanced parameter estimation and classification algorithms we expect to be able to classify the targets with high efficiency. This objective concerns the component of the classification problem that involves false alarm reduction.

3.2.1 Metric

The metric for this objective is the number of items-of-interest on the master dig list that can be correctly classified as non-munitions by each classification approach.

3.2.2 Data Requirements

Each demonstrator will prepare a prioritized dig list for the targets on the master anomaly list. IDA personnel will use their scoring algorithms to assess the results.

3.2.3 Success Criteria

The objective will be considered to be met if more than 30% of the non-munitions items can be correctly labeled as non-munitions while retaining all of the targets-of-interest on the dig list.

3.3 OBJECTIVE: SPECIFICATION OF NO-DIG THRESHOLD

In a retrospective analysis as will be performed in this demonstration, it is possible to tell the true classification capabilities of a classification procedure based solely on the prioritized dig list submitted by each demonstrator. In a real-world scenario, all targets may not be dug so the success of the approach will depend on the ability of an analyst to accurately specify their dig/no-dig threshold.

3.3.1 Metric

The probability of correct classification, P_{class}, and number of false alarms, N_{fa}, at the demonstrator-specified threshold are the metrics for this objective.

3.3.2 Data Requirements

Each demonstrator will prepare a ranked anomaly list with a dig/no-dig threshold indicated. IDA personnel will use their scoring algorithms to assess the results.

3.3.3 Success Criteria

The objective will be considered to be met if more than 30% of the non-munitions items can be correctly labeled as non-munitions while retaining all of the targets-of-interest at the demonstrator-specified threshold.

3.4 OBJECTIVE: MINIMIZE NUMBER OF ANOMALIES THAT CANNOT BE ANALYZED

Anomalies for which reliable parameters cannot be estimated cannot be classified by the classifier. These anomalies must be placed in the dig category and reduce the effectiveness of the classification process.

3.4.1 Metric

The number of anomalies for which reliable parameters cannot be estimated is the metric for this objective.

3.4.2 Data Requirements

Each demonstrator that estimates target parameters will provide a list of all parameters as part of their results submission along with a list of those anomalies for which parameters could not be reliable estimated.

3.4.3 Success Criteria

The objective will be considered to be met if reliable parameters can be estimated for > 90% of the anomalies on each sensor anomaly list.

3.5 OBJECTIVE: CORRECT ESTIMATION OF TARGET PARAMETERS

This objective involves the accuracy of the target parameters that are estimated in the first phase of the analysis. Successful classification is only possible if the input features are internally consistent. The obvious way to satisfy this condition is to estimate the various target parameters accurately.

3.5.1 Metric

Accuracy of estimation of target parameters is the metric for this objective.

3.5.2 Data Requirements

Each demonstrator that estimates target parameters will provide a list of all parameters as part of their results submission. IDA analysts will compare these estimated parameters to those measured during the intrusive investigation and determined via subsequent in-air measurements.

3.5.3 Success Criteria

The objective will be considered to be met if the estimated X, Y locations are within 15 cm (1σ), the estimated depths are within 10 cm (1σ), and the estimated size is within ± 20%.

4.0 SITE DESCRIPTION

The former Camp Beale is approximately 60,000 acre site located in Yuba and Nevada Counties, California. The demonstration was conducted in a 10 acre area that is located within the historical bombing Target 4 and the Proposed Toss Bomb target area. See the Program Office demonstration plan for more details on the site.

4.1 SITE SELECTION

This site was chosen as the next in a progression of increasingly more complex sites for demonstration of the classification process. The first site in the series, Camp Sibert, was open field survey conditions with only one target-of-interest and item "size" was an effective discriminant. At this site, there a wide mixture of munition types was anticipated and the site was partially wooded, adding another layer of complexity into the process. The survey areas consist of an "Open Area" on which the MetalMapper and Geonics EM61 collected data, and a "Portable Area" on which the Geonics EM61, BUD, TEMTADS and MPV collected data. There is an area of overlap.

4.2 SITE HISTORY

See the Program Office demonstration plan.

4.3 SITE GEOLOGY

See the Program Office demonstration plan.

4.4 MUNITIONS CONTAMINATION

See the Program Office demonstration plan.

5.0 TEST DESIGN

See the Program Office demonstration plan for a description of the test design for the overall project.

Sky Research/UBC-GIF processed data and delivered the following six digsheets for the Camp Beale Portable Area:

1) Cart EM-61 statistical: Discrimination achieved using features derived from the data amplitude versus decay feature space;
2) HandHeld Berkley UXO Discriminator (BUD) library match: A digsheet was produced based on matches to library items identified via test pit measurements and training data requests;
3) HandHeld (BUD) Support Vector Machine (SVM): We provided an alternative ranking of the HandHeld BUD based on SVM;
4) Man Portable Vector (MPV) library match: as in 2) but for the MPV data;
5) MPV SVM: as in 3) but for the MPV data.

Sky Research/UBC-GIF processed data and delivered the following eight digsheets for the Camp Beale Open Area:

1) Cart EM-61 statistical: Discrimination achieved using features derived from the data amplitude versus decay feature space;
2) CH2M Hill MetalMapper library match: A digsheet was produced based on matches to library items identified via test pit measurements and training data requests;
3) CH2M Hill MetalMapper SVM: We provided an alternative ranking of the CH2M Hill MetalMapper data based on SVM;
4) Parsons MetalMapper library match: as in 2) but for the Parsons MetalMapper data;
5) Parsons MetalMapper library match: as in 3) but for the Parsons MetalMapper data;
6) Prescreened CH2M Hill MetalMapper library match: same as 2) except 231 high confidence non-TOI items identified in EM-61 data were assumed non-TOI and not analyzed.
7) Prescreened CH2M Hill MetalMapper SVM: as in 6) with SVM methods used for classification.
8) Prescreened Parsons MetalMapper library match: as in 6) but for the Parsons MetalMapper data; and
9) Prescreened CH2M Hill MetalMapper SVM: as in 7) but for the Parsons MetalMapper data.

6.0 DATA ANALYSIS AND RESULTS

In this section we describe our data analysis procedures in general. A more detailed discussion specific to many of the datasets summarized here are provided in Appendix A-E

The Camp Beale demonstration site was divided up into two distinct areas differentiated by the sensors deployed (see Figure 1). A relatively open area was surveyed by the MetalMapper and a more challenging treed survey area was surveyed by portable sensors including the HandHeld BUD, MPV and TEMTADS 2x2x3. There was also an overlap area that was surveyed by both the portable systems and the MetalMapper to provide comparisons of all sensors for identical targets. EM61 data was first acquired over all survey areas and used to pick targets for cued investigations. Sky Research proposed to process all available datasets in order to test practical UXO classification methods over the complete set of EMI instruments deployed. The workflow was similar for all datasets and is outlined in Figure 2. Classification was performed by a different Sky analyst for each dataset with a strict firewall maintained throughout the analyses.

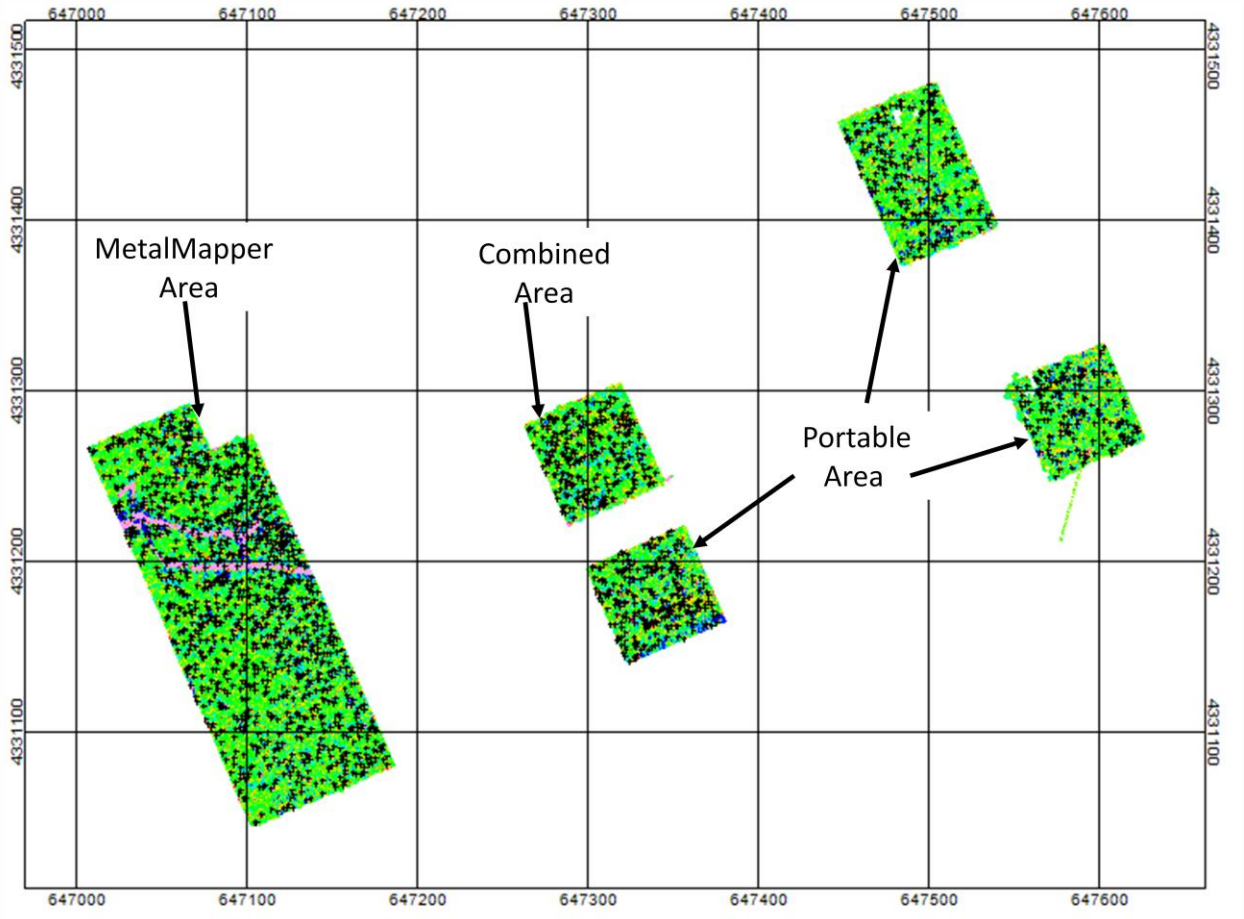


Figure 1: Full coverage EM61 data was collected and used to pick targets for cued investigation. Portable sensors (HandHeld BUD, MPV, TEMTADS 2x2) were deployed in a treed area not feasible for vehicular sensors and the MetalMapper was deployed in an open field area. The Combined Area was surveyed by all instruments.

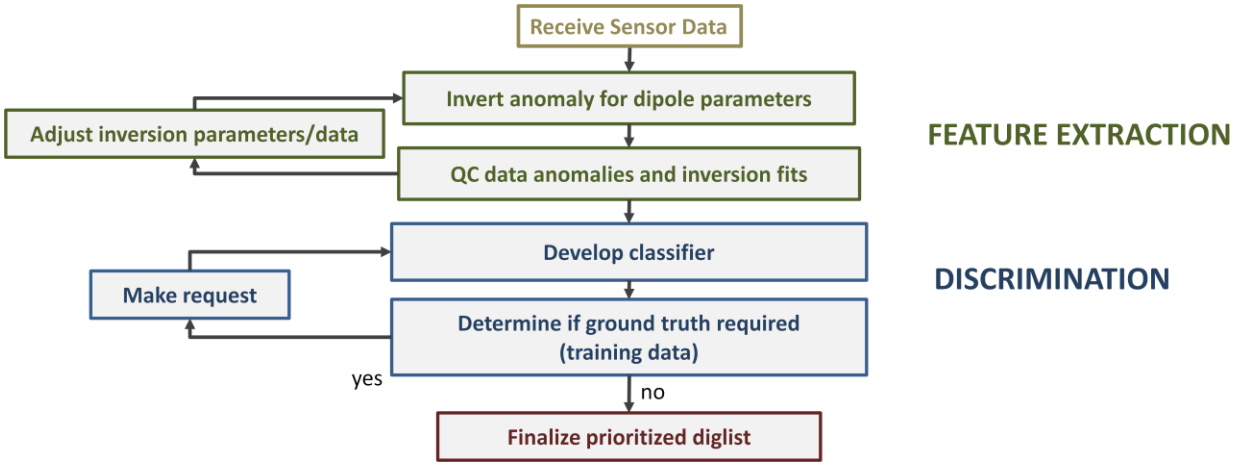


Figure 2: Overview of workflow applied to each dataset from Camp Beale.

6.1 Processing Portable Area Datasets

Sky generated diglists for all sensors deployed in the Portable area at Camp Beale. Photos of the portable sensors are illustrated in Figure 3. Processing a wide range of sensor data helps reinforce the sensor independent nature of the processing algorithms and classification methodologies. It also provides an opportunity to compare the relative performance of the various sensors as summarized in section 6.1.5.

Figure 3: Sensors deployed in the Portable area at Camp Beale included, from left to right, Geonics EM61, TEMTADS 2x2x3, HandHeld BUD, MPV.

6.1.1 Geonics EM61 Mk2 Cart Results

EM61 ordered diglists were generated using a decay parameter estimated from dipole model fits. The target size is poorly constrained in EM61 data because of uncertainties in the depth so the maximum predicted amplitude of the data at the first time channel is used as a proxy for target size. Data amplitude is useful for identifying large TOI in the early stages of digging. Using a data amplitude vs. polarizability decay feature space (see Figure 4) provides an initial improvement in classification performance relative to a strategy that relies on decay alone. The lack of separable, distinct clusters in the EM61 feature space illustrates the challenges in applying feature based classification methods with EM61 data.

The data amplitude/decay classifier does surprisingly well identifying larger TOI, with approximately 30% of clutter excavated in order to find all TOI. This is a marked reduction in the EM-61 false alarm

rate for previous demonstrations with comparable TOI classes (e.g. Camp Butner). We attribute this difference to the distributions of clutter: there is a reasonable separation between the distributions of TOI and non-TOI at Camp Beale (see Figure 25). At Camp Butner, a significant proportion of clutter was of comparable size (and anomaly amplitude) to 37mm items, and so identification of this TOI class was quite challenging. The lesson is that while good classification performance can sometimes be achieved with the EM-61, it is highly dependent on the relative sizes of TOI and non-TOI.

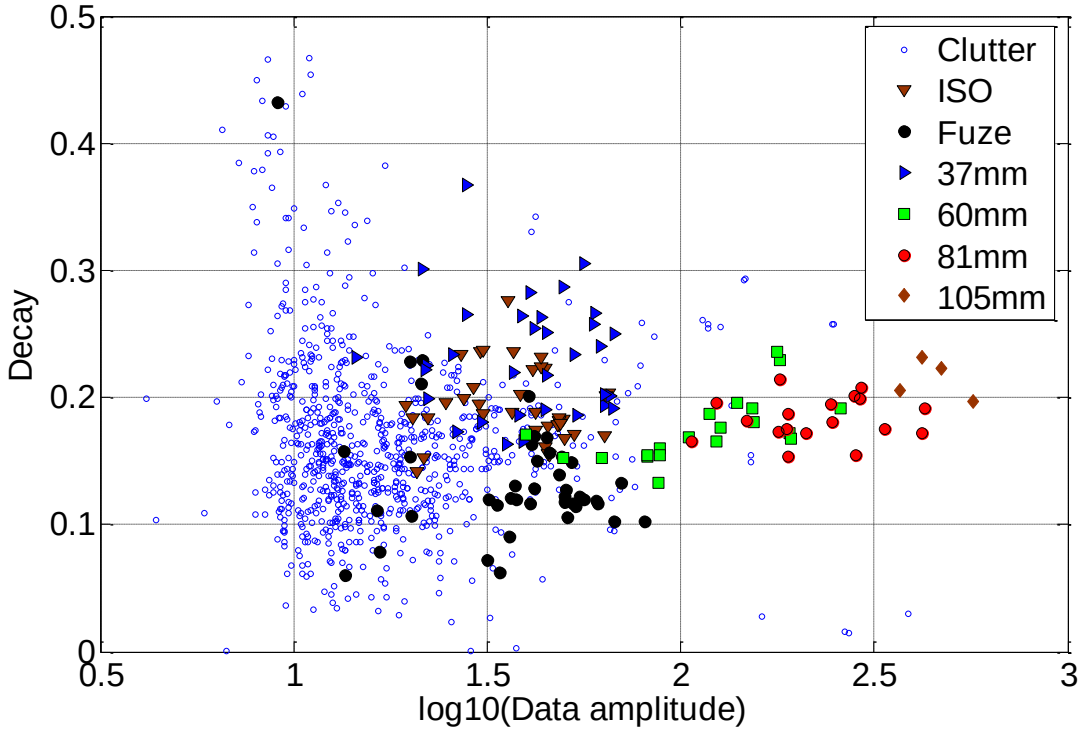


Figure 4: Feature space plot for the EM61 sensor in the Camp Beale Portable Area.

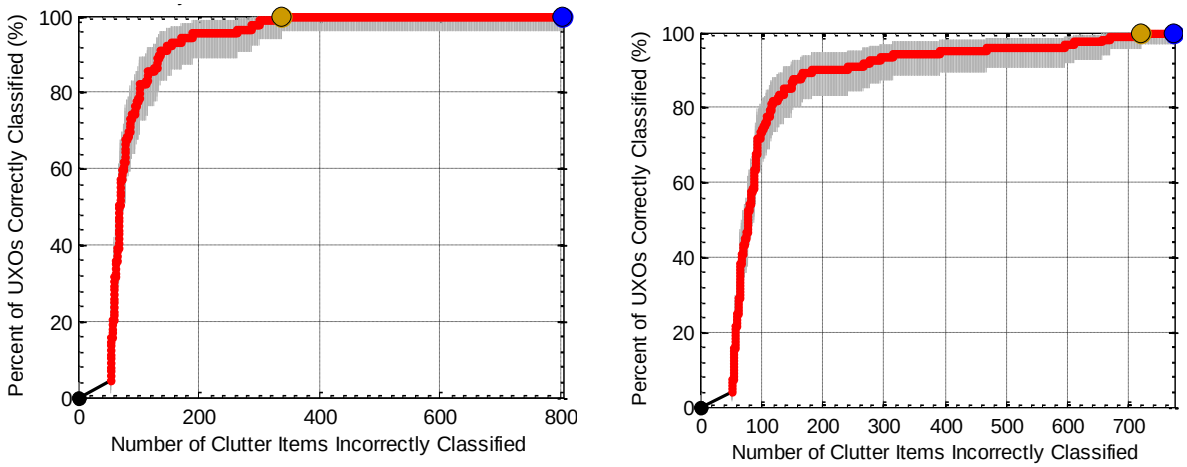


Figure 5: ROC curves for the EM61 sensor in the Camp Beale Portable Area. Results are shown both for fuzes considered clutter (left) and fuzes considered UXO (right).

6.1.2 HandHeld BUD Results

The BUD (Berkeley UXO discriminator) hand-held TEM system consists of three orthogonal transmitters and ten pairs of dB/dt induction receivers that measure gradient fields. Multiple BUD soundings were collected at each flagged anomaly. We chose to invert two soundings for each anomaly: one data set for the flag location and one selected with maximum ZZ response (Z receiver response for Z transmitter firing). The inversions were run with single and two-object models. Visual QC was performed in conjunction with several metrics of misfit, data SNR, and recovered polarizabilities. A diglist was generated by a UBC analyst using a library matching method to rank anomalies. A second Sky Research analyst generated an SVM diglist based on the inversions and training data requested by the UBC analyst. ROC curves are shown in Figure 7 for both the library matching and SVM diglists.

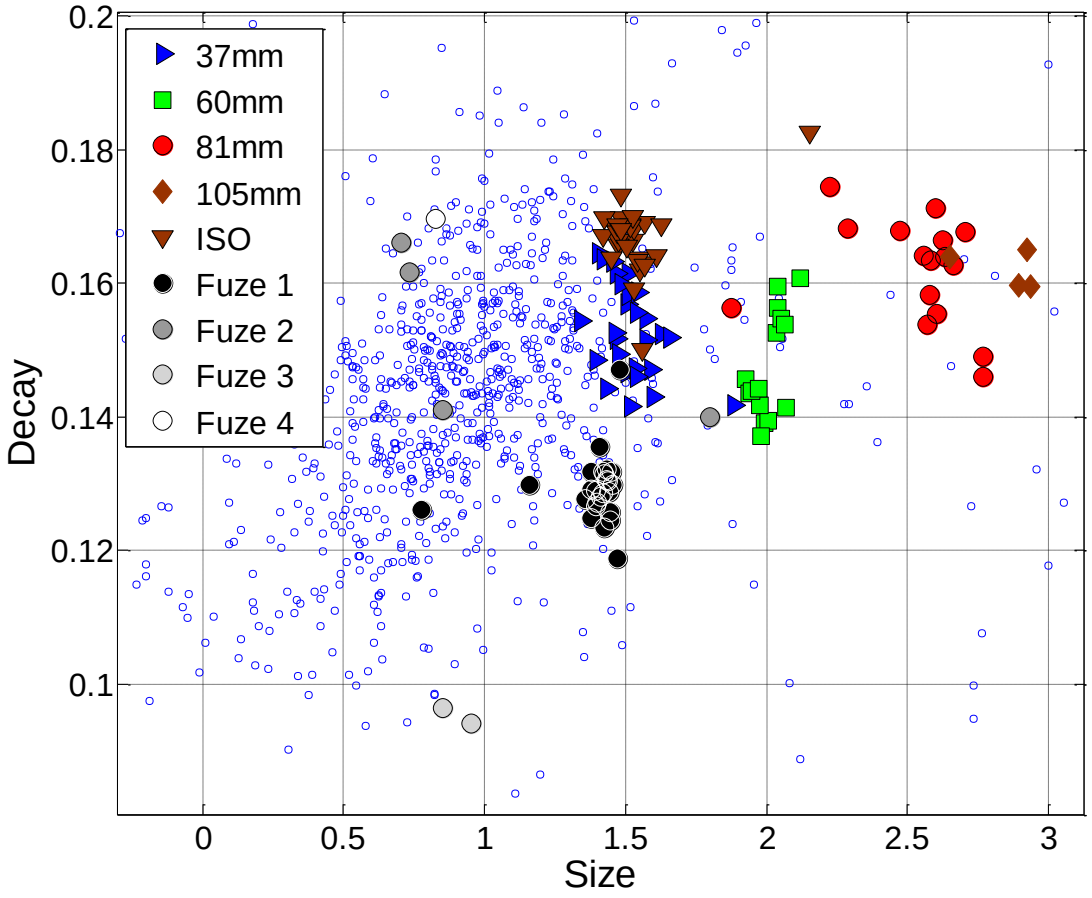


Figure 6: Feature space plot for the HandHeld BUD sensor in the Camp Beale Portable Area.

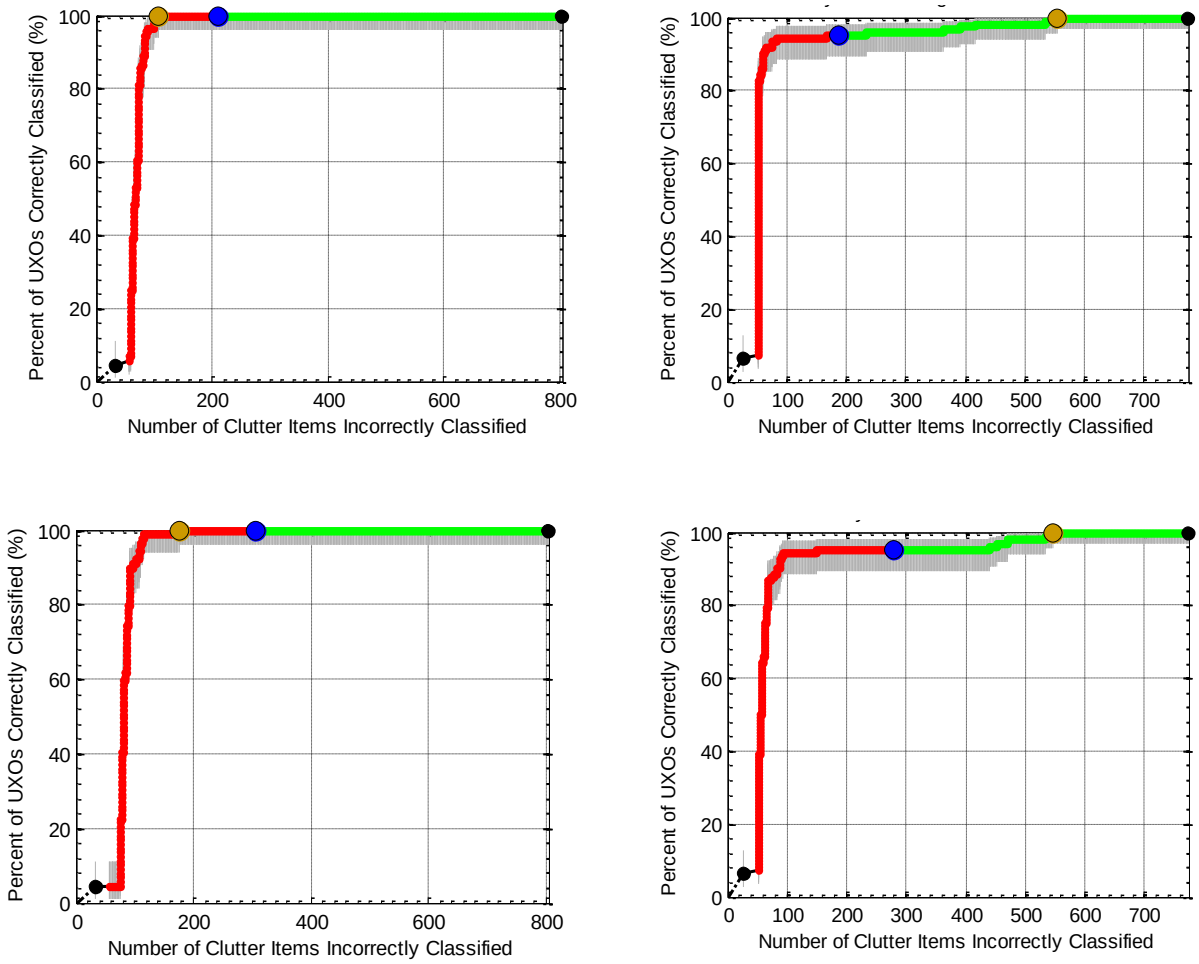


Figure 7: ROC curves for the HandHeld BUD. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right).

6.1.3 TEMTADS 2x2x3 Results

TEMTADS 2x2x3 data were inverted using both a single object and two object solvers. The SKY Research QC Tool was then used to review the one and two dipole fits and the model that was most similar to a known TOI (and that produced a good fit to the data) was selected . Three stages of ground truth information were requested. The initial ground truth request was comprised of 29 items designed to identify a number of suspicious looking items including fuzes. TOI contained in the initial ground truth request included two 37mm projectiles, four 81mm mortars, three fuzes and one ISO. A second round of ground truth contained 56 items designed to target uncertainty including poorly fitting models and items with suspicious (but not convincing) polarizabilities. The second batch of ground truth contained two more each of 37mm projectiles and 60mm mortars as well as four additional fuzes and four ISOs. Finally, a third set of ground truth was requested for six items targeting potential small fuzes. When all six items were revealed to be clutter, the analyst was confident that sufficient training data had been obtained to train the classifier.

Two SVM classifiers were then trained using the TOI identified via ground truth requests as well as feature vectors of 500 items that were judged to be least like the TOI and considered non-TOI. The first classifier used all polarizabilities (L1 up to the 88th time channel; L2 and L3 up to the 66th time channel). The second classifier used the total polarizability (L1+L2+L3 up to the 66th time channel). In creating the prioritized dig list, initial rankings were determined using all polarizabilities then the analyst manually determined a point in the dig list where anomalies would be ranked based on total polarizability. This point was chosen by observing where the high polarizability (obtained by visual review) anomalies occurred in the dig list. Some low ranked anomalies were moved up into a "can't decide" class and dug as potential TOI. The stop dig point was also determined based on manual review and was comprised of 228 items plus 31 anomalies declared as "can't decide". Based on the ground truth results obtained up to the submitted stop dog point, a further 24 items were dug (none of which were TOI). ROC curves for the submitted diglist are shown in Figure 9.

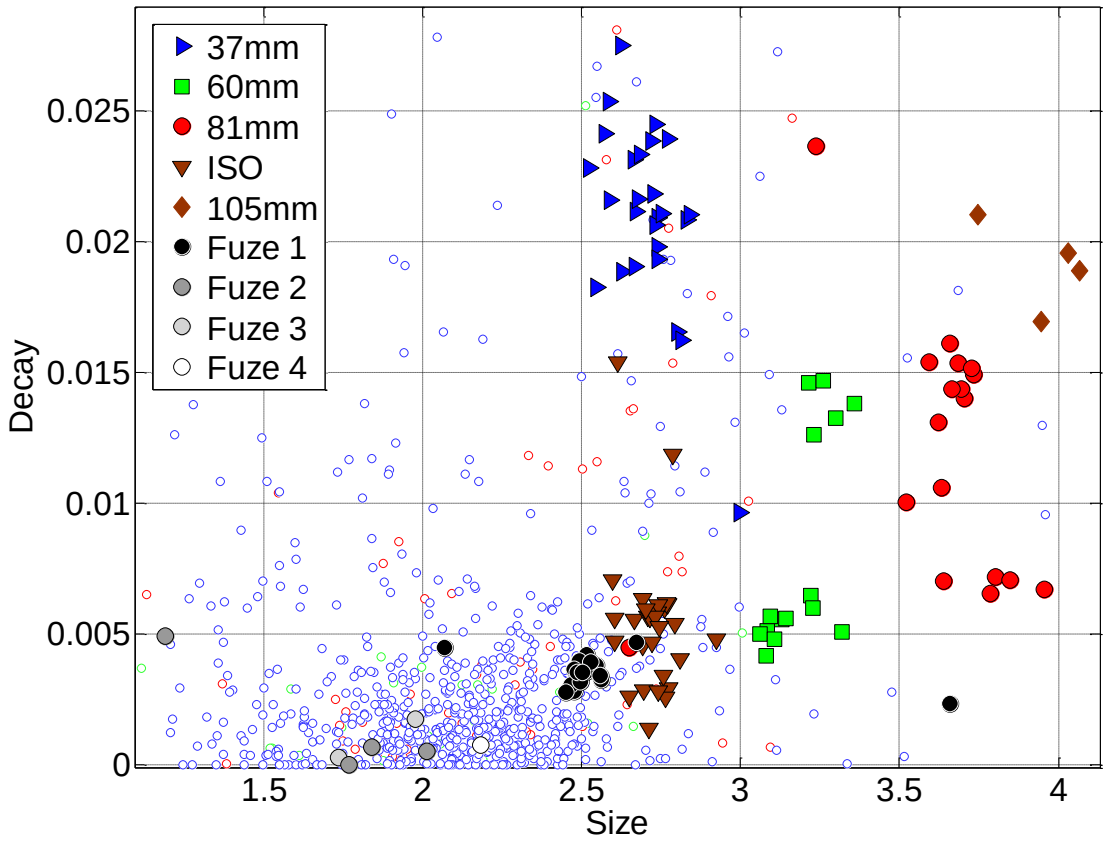


Figure 8: Feature space plot for the TEMTADS 2x2x3 sensor in the Camp Beale Portable Area.

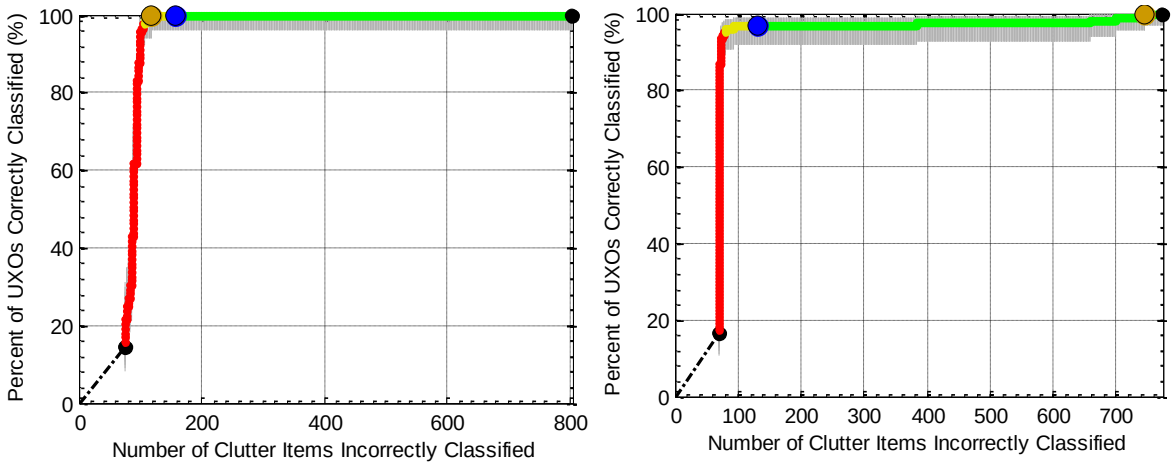


Figure 9: ROC curves for the TEMTADS 2x2x3 sensor. Results are shown both for fuzes considered clutter (left) and fuzes considered UXO (right).

6.1.4 MPV Results

Data from the Man Portable Vector (MPV) sensor was processed by a Sky analyst from a simplified production level scenario for comparison with results obtained via a more thorough R&D level analysis performed by an independent Sky analyst. The training data used were the same as those chosen for the R&D analysis. Only a single object inversion was performed for each anomaly and an expedited review of fits was performed using the Sky Research QC tool. All fits were reviewed in a single day, with less time devoted to investigation of data peculiarities than would normally occur in a typical R&D level analysis

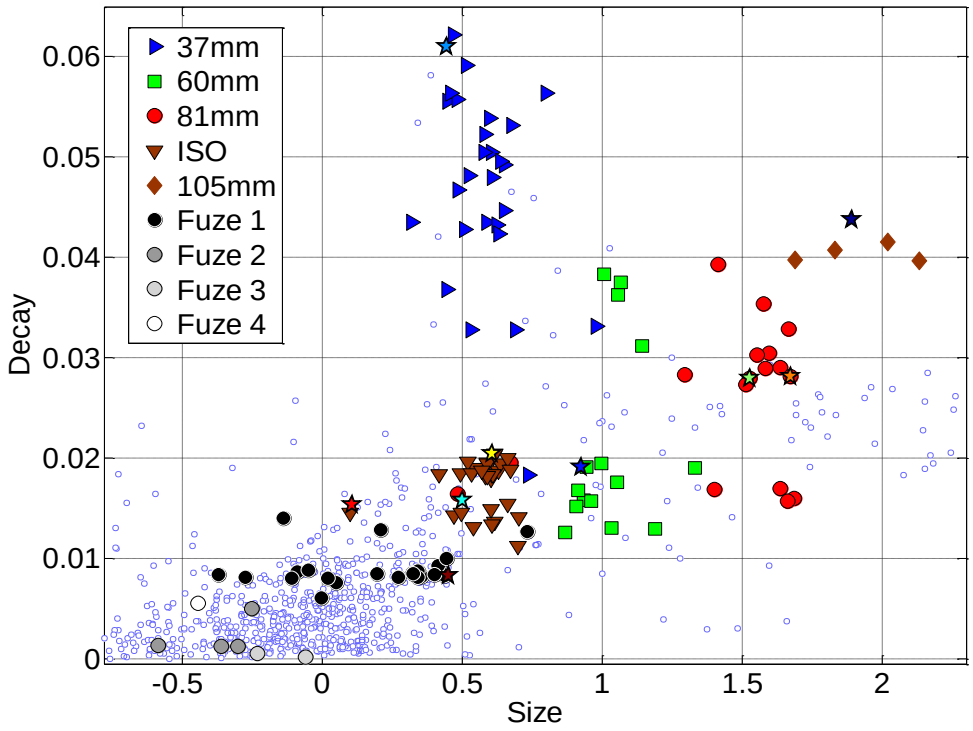


Figure 10: Feature space plot for the MPV sensor in the Camp Beale Portable Area.

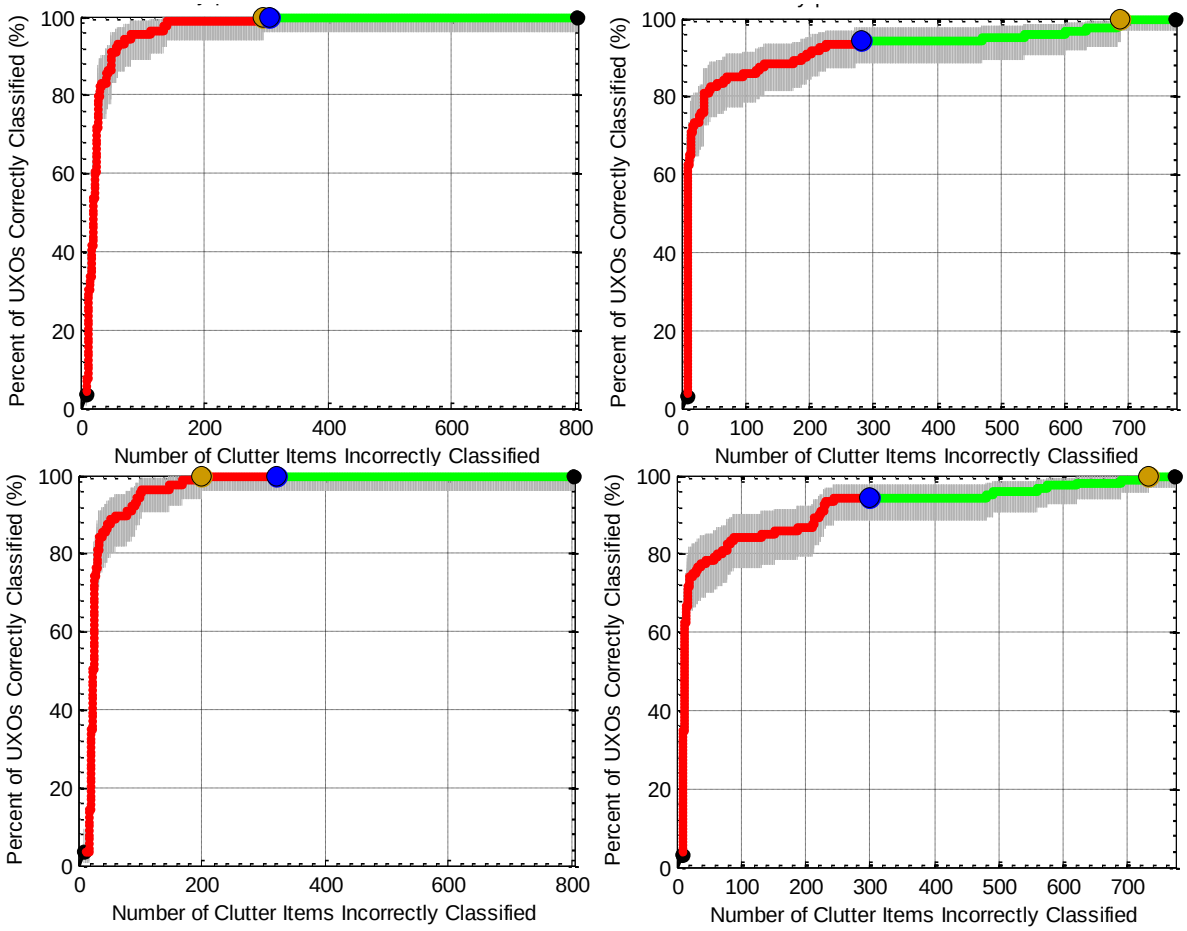


Figure 11: ROC curves for the MPV. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right).

6.1.5 Portable Area Summary

Performance for portable sensors was remarkably consistent. For the case with fuzes considered clutter, all sensors were able to identify all UXO. The amount of scrap dug varied from 19.6% for the TEMTADS

2x2x3 to 100% for the EM61. The SVM method digs a higher percentage of scrap than the polarizability match methods because it uses an automatically determined stop dig point, which tends to be more conservative than a stop dig point visually selected by an analyst.

For the case where fuzes are considered UXO, all methods missed some of the difficult fuzes. Successful detection of the small (5cm and smaller) fuzes would require modifying the method used in this demonstration to search for "onesies and twosies" as these were all unique target types with distinct polarizabilities and features that significantly overlapped with clutter items in feature space.

Sensor	Classification Method	Percentage of UXO Identified at Stop Dig Point	Percentage of Scrap Dug at Stop Dig Point	Missed TOI
MPV (SOI only)	Support Vector Machine	100	40.1	0
BUD	Support Vector Machine	100	37.9	0
TEMTADS 2x2	Support Vector Machine	100	19.5	0
MPV (SOI only)	Polarizability Match	100	37.9	0
BUD	Polarizability Match	100	26.2	0
EM61	Data Amplitude/Decay	100	100	0

Table 4: Summary of performance for all sensors in the Portable area when fuzes are considered clutter.

Sensor	Classification Method	Percentage of UXO Identified at Stop Dig Point	Percentage of Scrap Dug at Stop Dig Point	Missed TOI
MPV (SOI only)	Support Vector Machine	94.2	38.8	7
BUD	Support Vector Machine	95.0	36.3	6
TEMTADS 2x2	Support Vector Machine	96.7	16.9	4
MPV (SOI only)	Polarizability Match	94.2	36.4	7
BUD	Polarizability Match	95.0	24.2	6
EM61	Data Amplitude/Decay	100	100	0

Table 5: Summary of performance for all sensors in the Portable area when fuzes are considered TOI.

6.2 Processing Open Area Datasets

The Open area Camp Beale was surveyed by both EM61 and MetalMapper sensors. The EM61 data was acquired in full coverage mode in order to pick anomalies for subsequent follow up with the MetalMapper in cued investigation survey mode. Cued MetalMapper data was acquired by CH2M Hill and Parsons over an identical set of targets using the same instrument. Two different Sky analysts created ordered diglists for each dataset to ensure a firewall was maintained for training data requests and groundtruth.

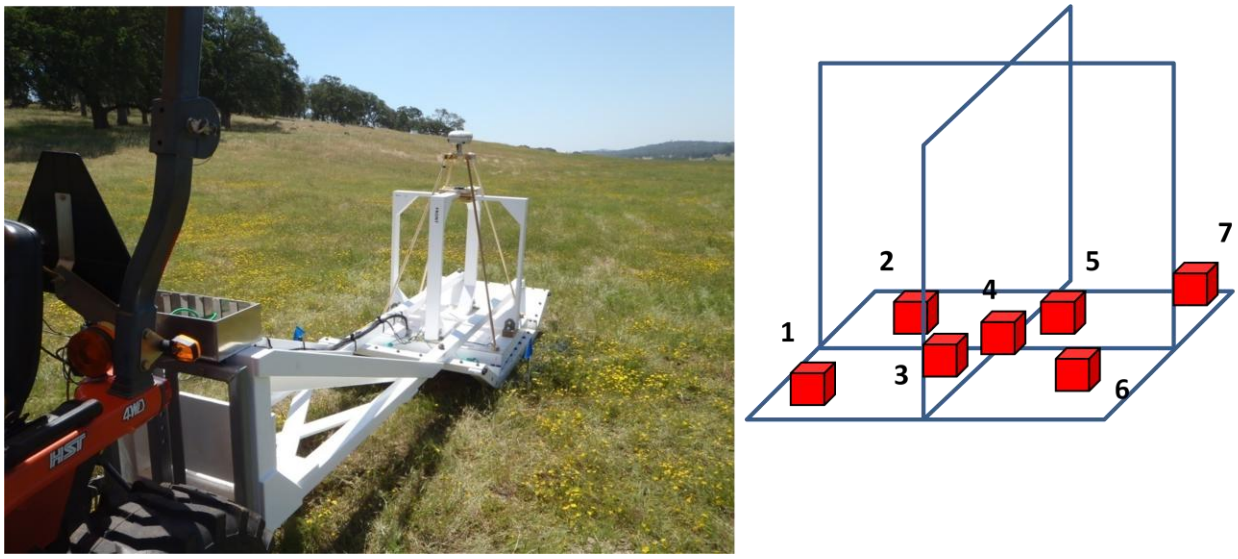


Figure 12: MetalMapper sensor collecting data at Camp Beale (left). Sensor consists of three orthogonal transmitters and seven multi-static receiver cubes (on right).

6.2.1 Geonics EM61 Mk2 Cart Results

Diglists in the Open area were generated using the same approach discussed in section 6.1.1 for the Portable area. A size versus decay feature space is shown in Figure 13 and the ROC curves for the submitted diglists are shown in Figure 14. Similar to the findings for the EM61 in the portable area, the reduction of digs in order to identify all TOI is not nearly as substantial as results obtained with the cued sensor data because the classification capabilities of data acquired with the EM61 sensor is limited when compared with the multiple target illuminations and longer time windows available using the purpose built sensors.

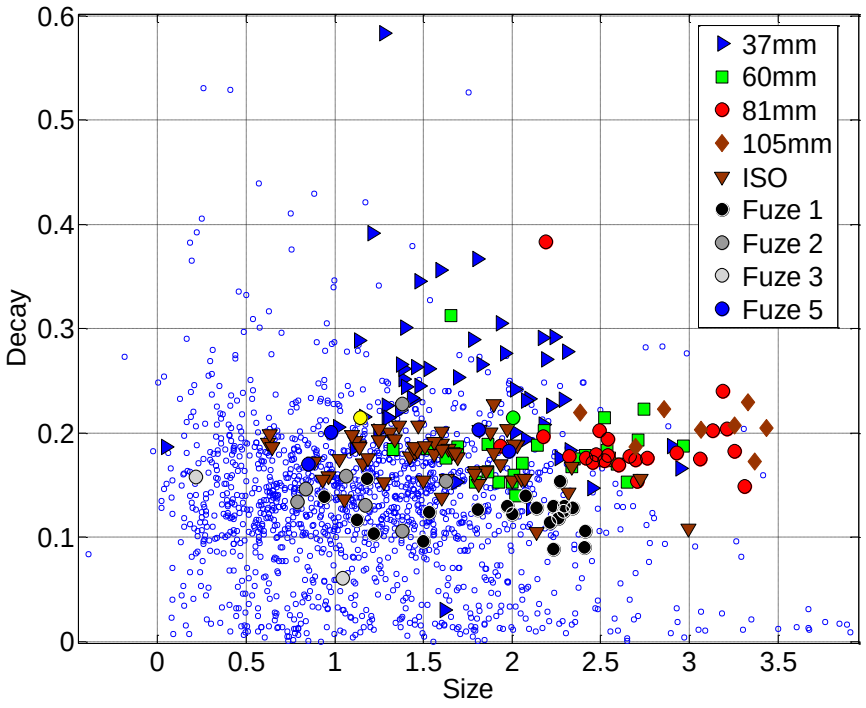


Figure 13: Feature space for the EM61 sensor in the Camp Beale Open Area.

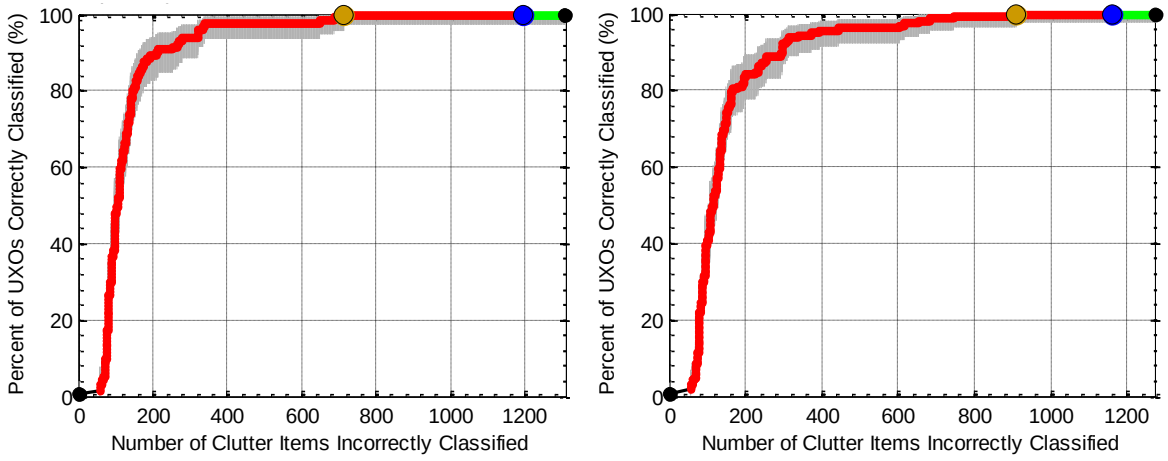


Figure 14: ROC curves for the EM61 sensor in the Camp Beale Open Area. Results are shown both for fuzes considered clutter (left) and fuzes considered UXO (right).

6.2.2 CH2M Hill MetalMapper Results

CH2M Hill MetalMapper data were inverted using both with a single object and two object inversions. The SKY Research QC Tool was then used to review the one and two dipole fits and identify the response most similar to a known TOI (and that produced a good fit to the data). Training data requests included

TOI with poorly resolved secondary polarizabilities as shown in Figure 15. As a result, a conservative approach was undertaken when formulating a diglist whereby all three polarizabilities were initially used to determine the dig order. After all the high confidence TOI were identified, the diglist was based on the primary polarizability only in order to identify TOI with poorly recovered secondary polarizabilities. Small fuses were also discovered in training data and attempts were made to identify those items before the stop dig point.

A size versus decay feature space is shown in Figure 16 and the ROC curves for the library match and SVM based classification methods.

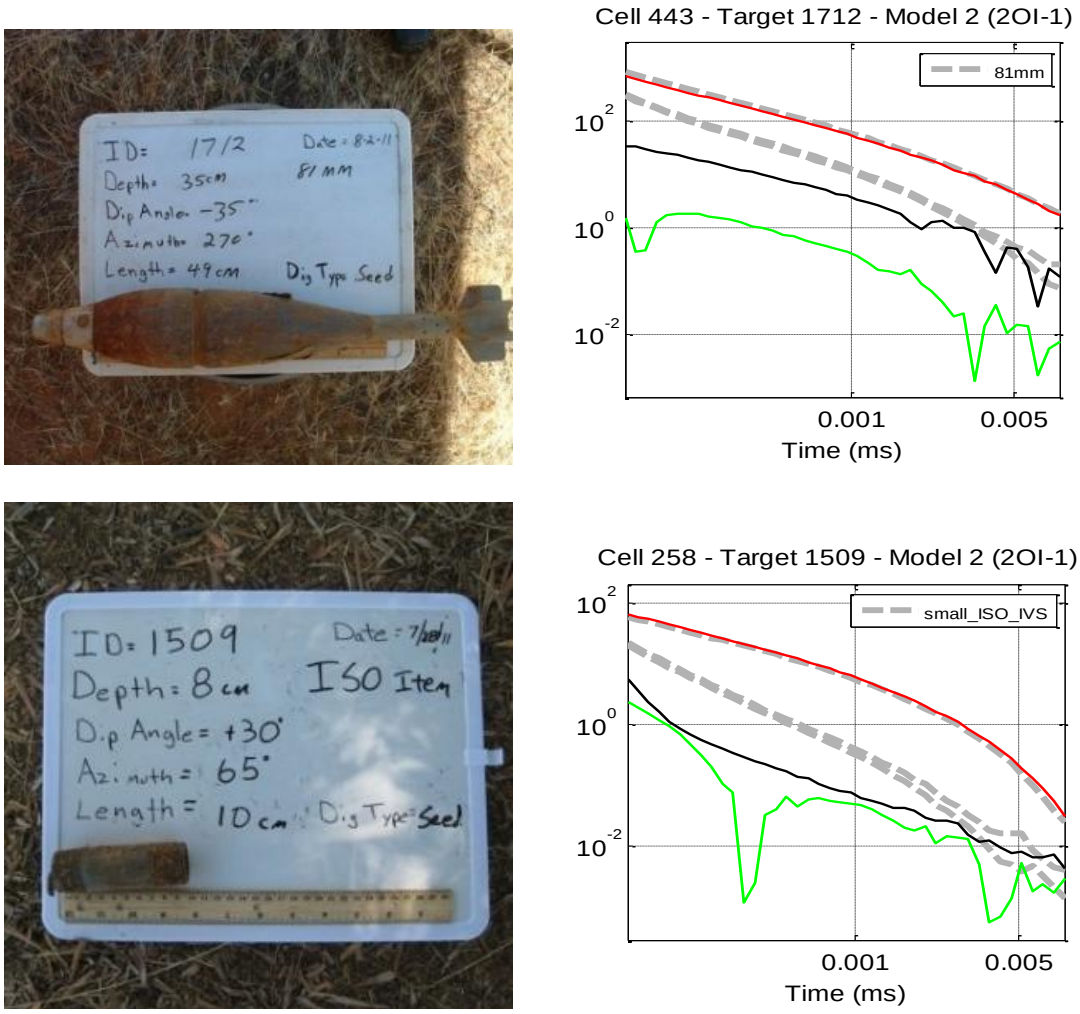


Figure 15: Training data for both TOI 1712 and 1509 both exhibited excellent matches to the primary polarizability of a library item but poorly constrained secondary polarizabilities.

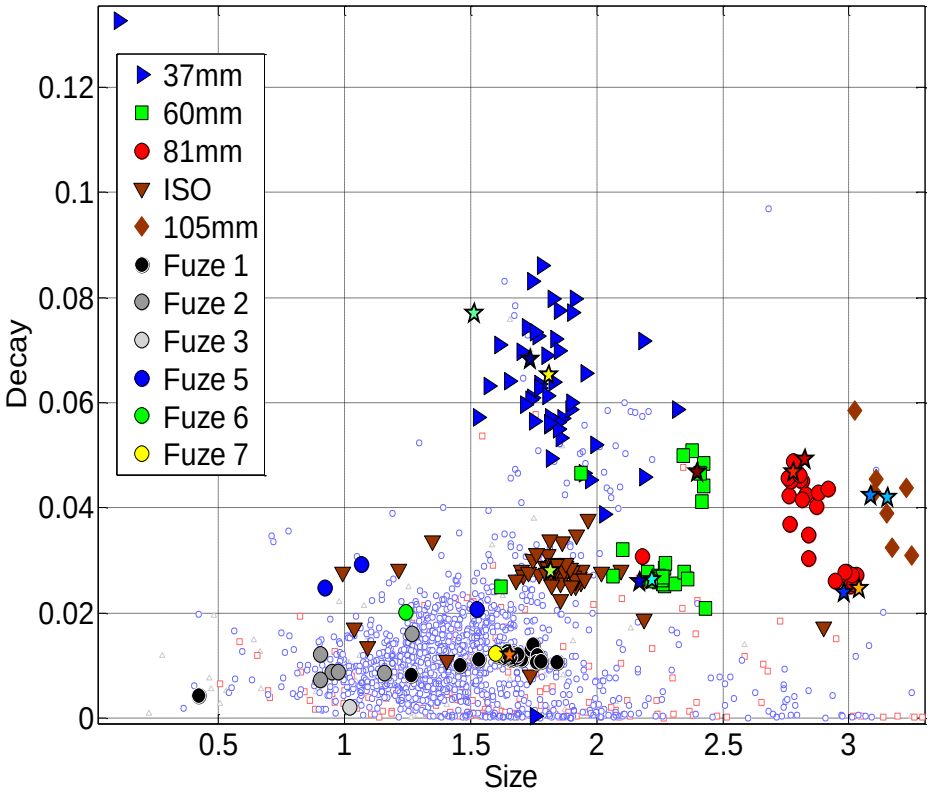


Figure 16: Feature space for the CH2M Hill MetalMapper sensor in the Camp Beale Open Area.

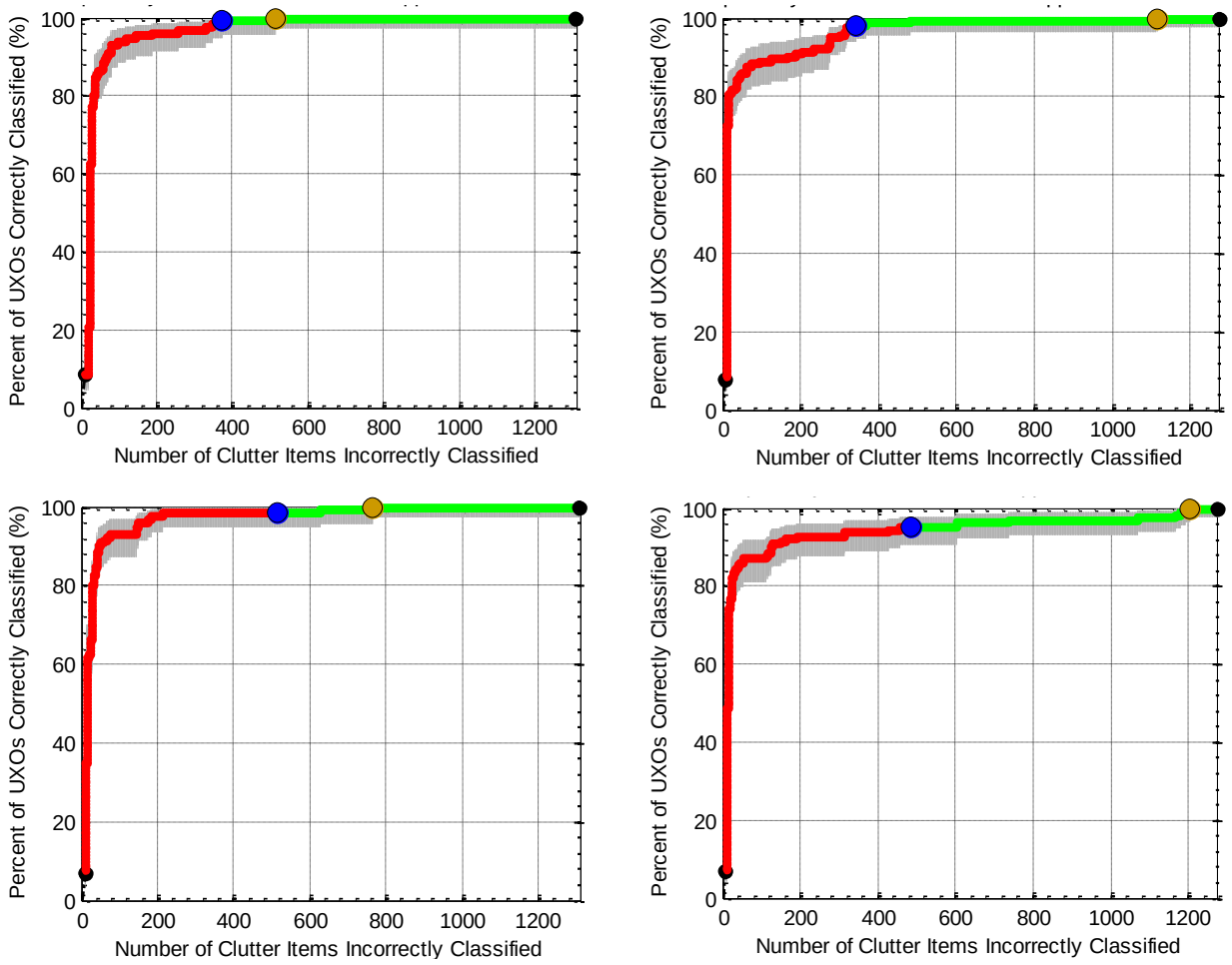


Figure 17: ROC curves for the CH2M Hill MetalMapper. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right).

6.2.3 Parsons MetalMapper Results

In contrast to the CH2M Hill case, training data requests did not include TOI with very poorly constrained secondary polarizabilities. This provided a somewhat overestimated sense of the quality of the Parsons data, leading to an over-aggressive dig list which relied on all three polarizabilities for all anomalies. Switching to matching on total polarizability would have resulted in the two missed ISO TOI being found significantly earlier. Ground truth from training data also did not include any small fuzes, hence the relatively poorer performance in the case of "fuzes as UXO". The excellent-looking partial ROC curve also contributed to a false sense of confidence in the data. The size decay feature space generated from the Parsons MetalMapper data is shown in

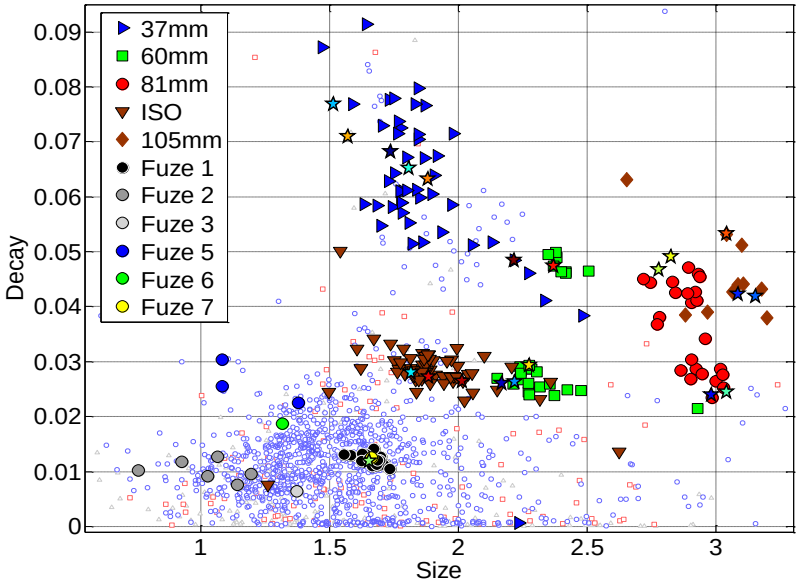


Figure 18: Feature space for the Parsons MetalMapper sensor in the Camp Beale Open Area.

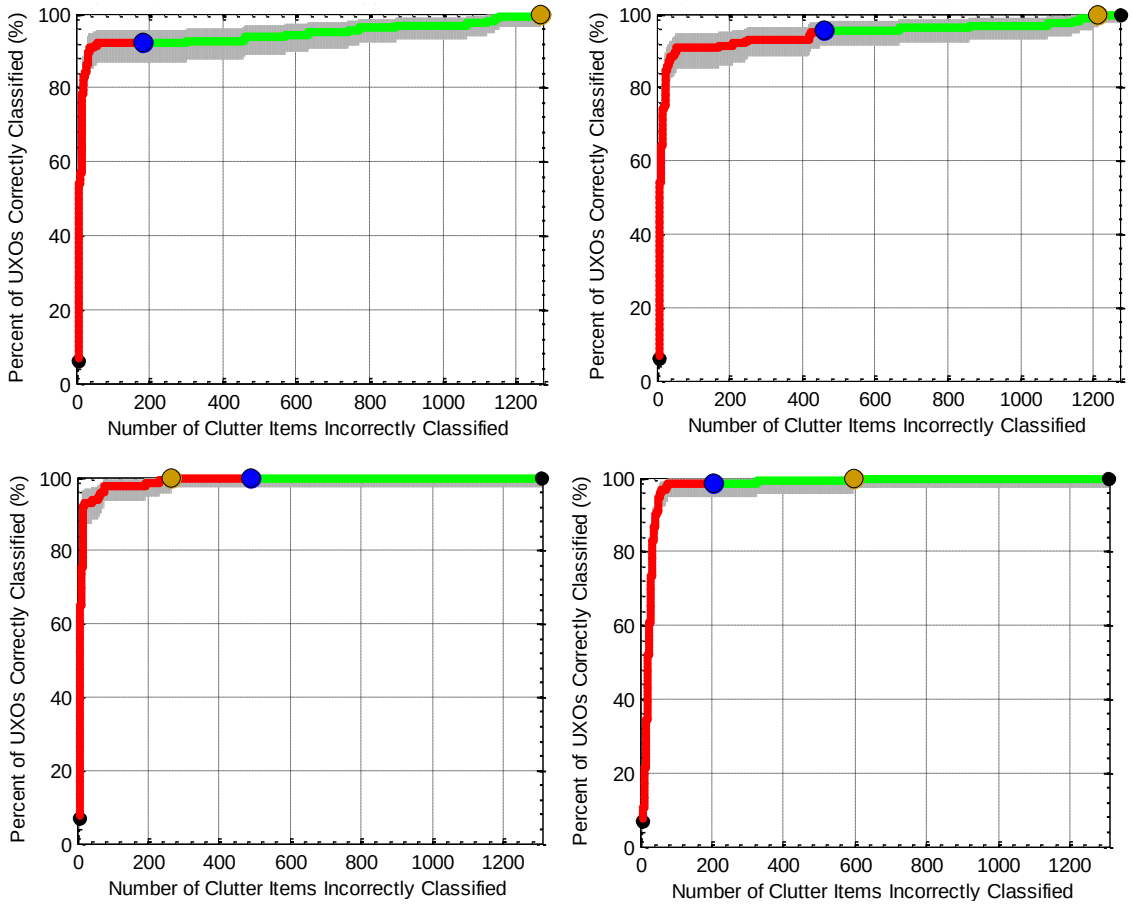


Figure 19: ROC curves for the Parsons MetalMapper. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results for the SVM diglist are on the bottom row (fuzes as clutter on left, fuzes as TOI on right).

6.2.4 Prescreened MetalMapper Results

Prescreened diglists were also submitted for both the CH2M Hill and Parsons MetalMapper datasets. A small number (231) of EM-61 anomalies were pre-screened on the basis of their location in data amplitude/decay space. Pre-screened test vectors were farthest from TOI training vectors in this two-dimensional space (see Figure 26). The first data request assumed that these targets would not be reacquired with the MetalMapper. Despite a very conservative cut-off, the set of pre-screened targets included a seed item (target 2277) and a number of small fuzes. At the 2010 Camp Butner demonstration, pre-screening also mislabeled TOI items. We therefore conclude that pre-screening with the EM-61 is only feasible at sites where targets of interest are large (105 mm, 4.2") and easily distinguished from native clutter. If smaller TOI are encountered, then all detected targets should be reacquired with a cued sensor. In future demonstrations we will not pursue pre-screening techniques. Results are shown in Figure 20 and Figure 21 for the diglists submitted. Again two different classification methodologies were attempted and diglists were generated for both polarizability match and SVM methods.

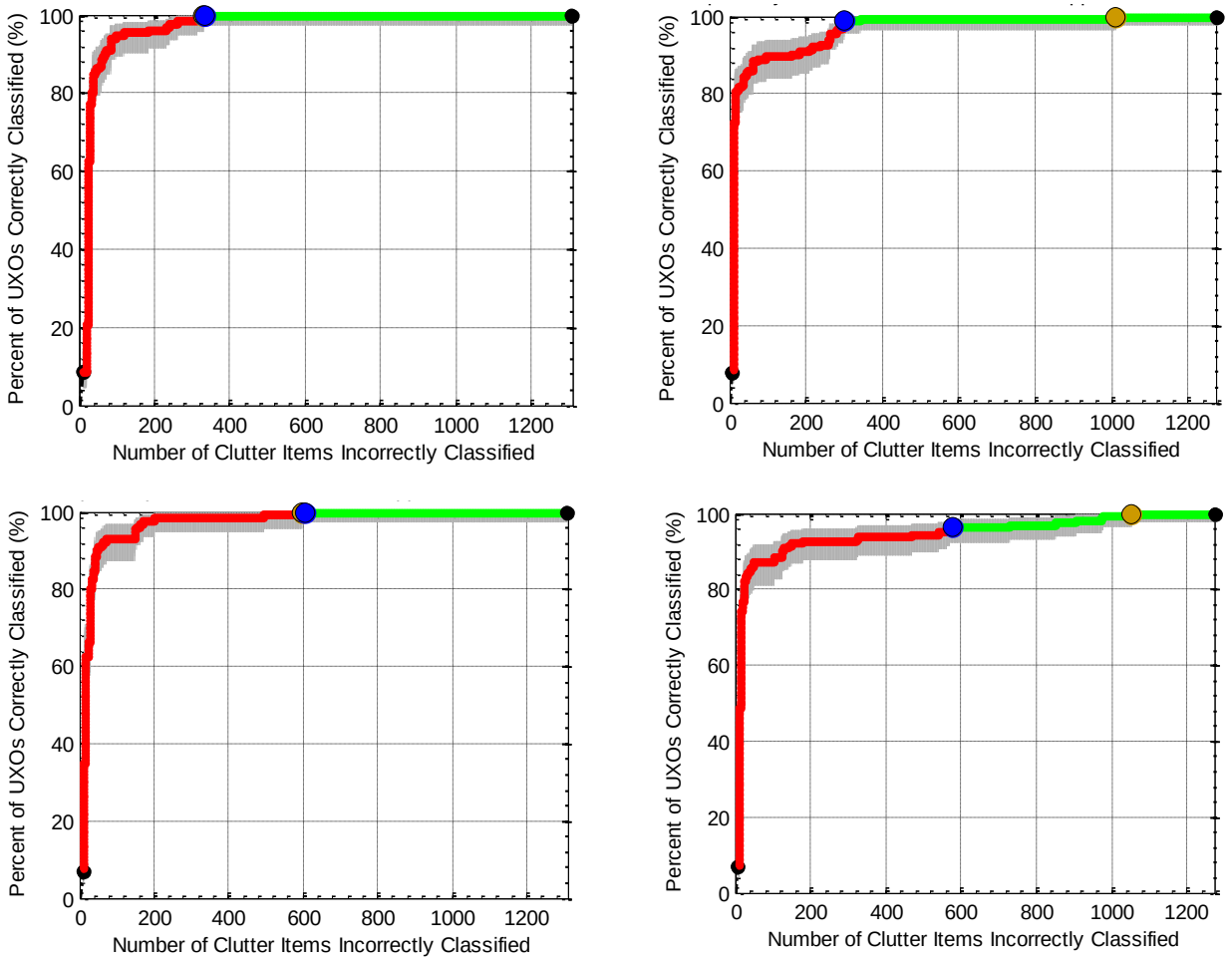


Figure 20: ROC curves for the prescreened CH2M Hill MetalMapper data. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right).

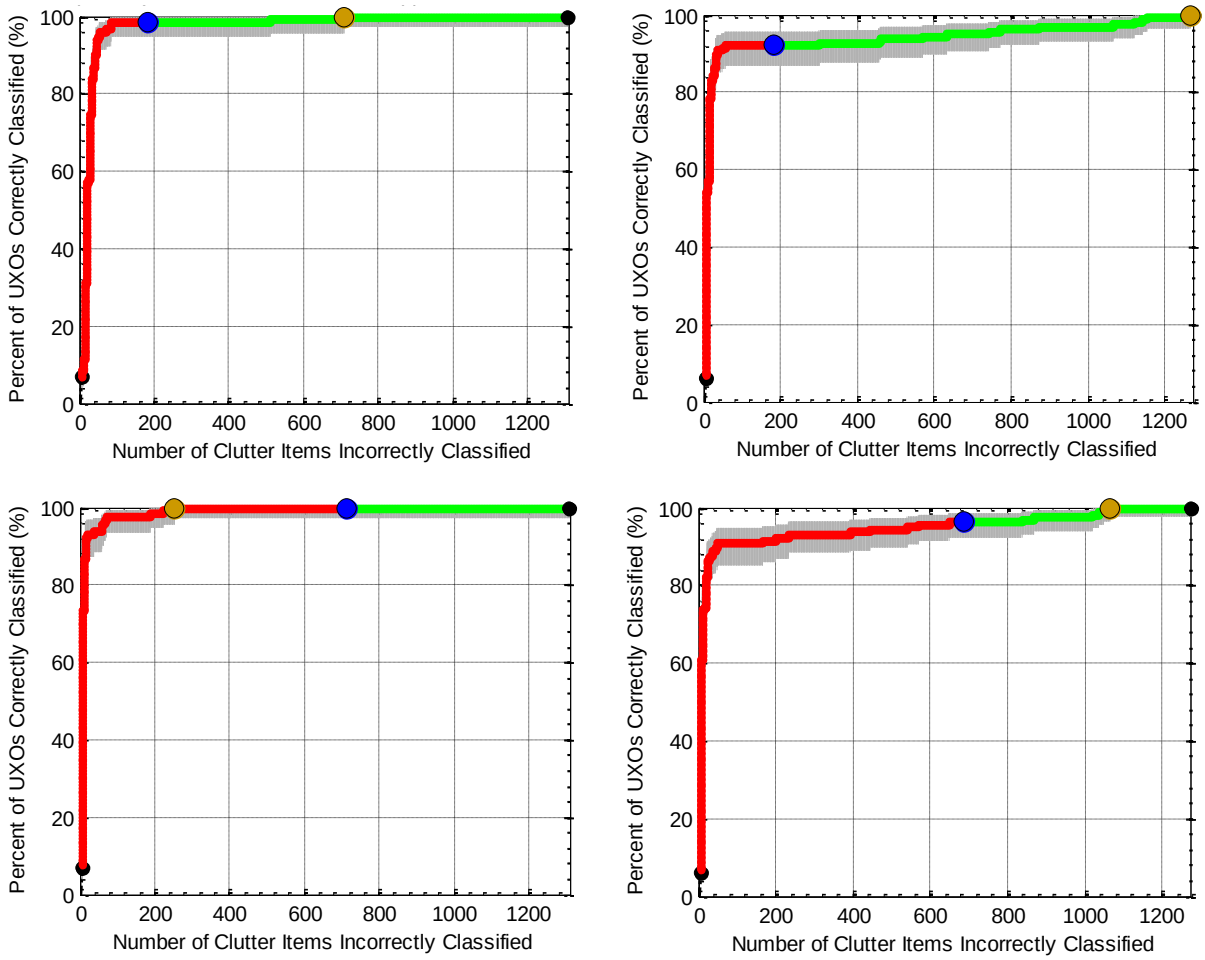


Figure 21: ROC curves for the prescreened Parsons MetalMapper data. Results are shown for two classification methods. Polarizability match diglist results are show in the top row (fuzes as clutter on the left, fuzes as TOI on right). Results obtained for the SVM diglist are illustrated in the bottom row (fuzes as clutter on the left, fuzes as TOI on right).

6.2.5 Open Area Summary

Performance of sensors in the Open Area at Camp Beale is summarized in Table 6 and Table 7. For the case with fuzes considered clutter, all sensors were able to identify all UXO. In cases where all TOI were identified, the amount of scrap dug varied from 24.1% for the prescreened MetalMapper data to 89% for the EM61. The SVM method digs a higher percentage of scrap than the polarizability match methods because it uses an automatically determined stop dig point, which tends to be more conservative than a stop dig point visually selected by an analyst. A number of methods missed one or two TOI. A detailed analysis of decisions leading to the missed TOIs is included in the Appendices.

For the case where fuzes are considered UXO, all methods missed some of the difficult fuzes. Successful detection of the small (5cm and smaller) fuzes would require modifying the method used in this demonstration to search for "onesies and twosies" as these were all unique target types with distinct polarizabilities and features that significantly overlapped with clutter items in feature space.

Sensor	Discrimination Method	Percentage of UXO Identified at Stop Dig Point	Percentage of Scrap Dug at Stop Dig Point	Missed TOI
EM61	Data Amplitude/Decay	100	89.0	0
MetalMapper (CH2M Hill)	Polarizability Match	99.2	27.4	1
MetalMapper (CH2M Hill)	Support Vector Machine	98.5	38.9	2
MetalMapper (Parsons)	Support Vector Machine	100	37.0	0
MetalMapper (Parsons)	Polarizability Match	98.5	16.3	2
Prescreened MetalMapper (CH2M Hill)	Polarizability Match	100	24.1	0
Prescreened MetalMapper (CH2M Hill)	Support Vector Machine	100	46.1	0
Prescreened MetalMapper (Parsons)	Support Vector Machine	100	55.0	0
Prescreened MetalMapper (Parsons)	Polarizability Match	98.5	13.5	2

Table 6: Summary of performance for all sensors in the Open area when fuzes are considered clutter.

Sensor	Discrimination Method	Percentage of UXO Identified at Stop Dig Point	Percentage of Scrap Dug at Stop Dig Point	Missed TOI
EM61	Data Amplitude/Decay	100	88.7	0
MetalMapper (CH2M Hill)	Polarizability Match	98.2	27.5	3
MetalMapper (CH2M Hill)	Support Vector Machine	95.1	39.4	8
MetalMapper (Parsons)	Support Vector Machine	95.7	37.6	7
MetalMapper (Parsons)	Polarizability Match	92.0	17.1	13
Prescreened MetalMapper (CH2M Hill)	Polarizability Match	98.8	24.3	2
Prescreened MetalMapper (CH2M Hill)	Support Vector Machine	96.3	46.5	6
Prescreened MetalMapper (Parsons)	Support Vector Machine	96.3	55.5	6
Prescreened MetalMapper (Parsons)	Polarizability Match	92.0	14.3	13

Table 7: Summary of performance for all sensors in the Open area when fuzes are considered TOI.

7.0 MANAGEMENT AND STAFFING

A flow chart showing the managerial hierarchy and the relationship between the principal investigator (PI) and other personnel is shown in Figure 22.

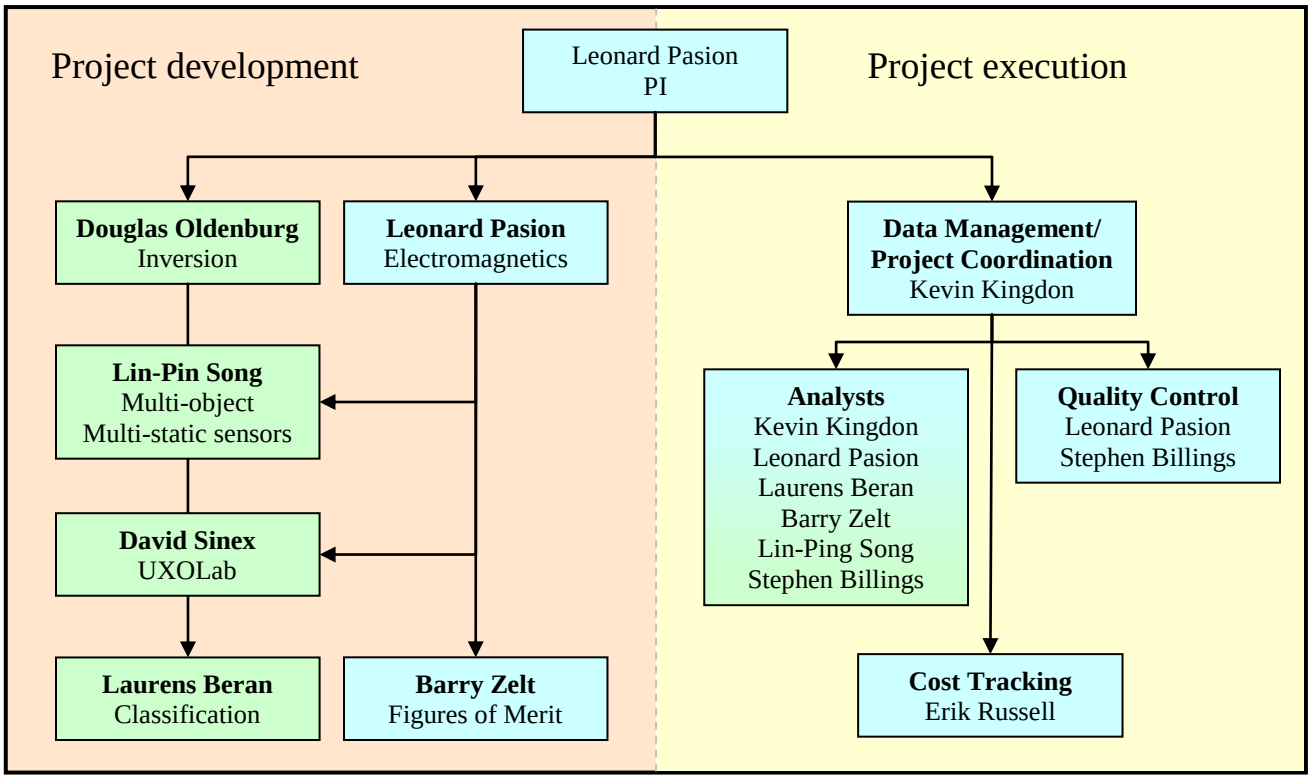


Figure 22. Project management hierarchy showing Sky Research personnel in blue and UBC-GIF personnel in green. The hierarchy is split between the development and execution components.

8.0 REFERENCES

T. Bell, B. Barrow, J. Miller, and D. Keiswetter. Time and frequency domain electromagnetic induction signatures of unexploded ordnance. *Subsurface Sensing Technologies and Applications*,2:153-175, 2001.

S. D. Billings, L. R. Pasion, L. Beran, N. Lhomme, L. Song, D. W. Oldenburg, K. Kingdon, D. Sinex, and J. Jacobson. Unexploded ordnance discrimination using magnetic and electromagnetic sensors: Case study from a former military site. *Geophysics*, 75:B103-B114, 2010.

ESTCP. 2010 ESTCP demonstration study, former Camp Butner. Technical report, 2010.

X. Liao and L. Carin. Migratory logistic regression for learning concept drift between two data sets with application to UXO sensing. *IEEE Trans. Geosci. Remote Sensing*, 47:1454 -1466, 2009.

L. R. Pasion, S. D. Billings, D. W. Oldenburg, and S.Walker. Application of a library-based method to time domain electromagnetic data for the identification of unexploded ordnance. *Journal of Applied Geophysics*, 61:279-291, 2007.

F. Shubitidze, K. O'Neill, S. A. Haider, K. Sun, and K. D. Paulsen. Application of the Method of Auxiliary Sources to the Wide-Band Electromagnetic Induction Problem. *IEEE Trans. Geosci. Remote Sensing*, 40:928-942, 2002.

S. L. Tantum, Y. Li, and L. M. Collins. Bayesian mitigation of sensor position errors to improve unexploded ordnance detection. *IEEE Geosci. Remote Sensing Letters*, 5:103{107, 2008.

G. F. West and J. C. Macnae. *Electromagnetic methods in applied geophysics*, chapter Physics of the electromagnetic exploration method, pages 5-45. SEG, 1991.

Y. Zhang, X. Liao, and L. Carin. Detection of buried targets via active selection of labeled data: Application to sensing subsurface UXO. *IEEE Trans. Geosci. Remote Sensing*, 42:2535-2543, 2004.

APPENDICES

Appendix A: EM61 FEATURE EXTRACTION AND CLASSIFICATION

To generate ordered diglists from Camp Beale EM-61 data sets, we used the *decay* parameter estimated from dipole model fits

$$\boldsymbol{decay} = \frac{\boldsymbol{L_{tot(t_4)}}}{\boldsymbol{L_{tot(t_1)}}} \tag{12}$$

with L_{tot} the total polarizability. Because target size is poorly constrained by EM-61 data (owing to poor constraints on depth), we do not use a size parameter derived from the dipole model. We instead use the maximum predicted amplitude of the data at the first time channel as a proxy for target size. Data amplitude is useful for identifying large TOI (e.g. 105 mm) in the early stages of digging, and so training in a data amplitude/decay feature space can provide some initial improvement in classification performance relative to a strategy that relies on decay alone.

Figure 23 shows the dig order for a probabilistic neural network (PNN) classifier applied to the Beale open field test data. This classifier was trained in a two-dimensional data amplitude/decay feature space.

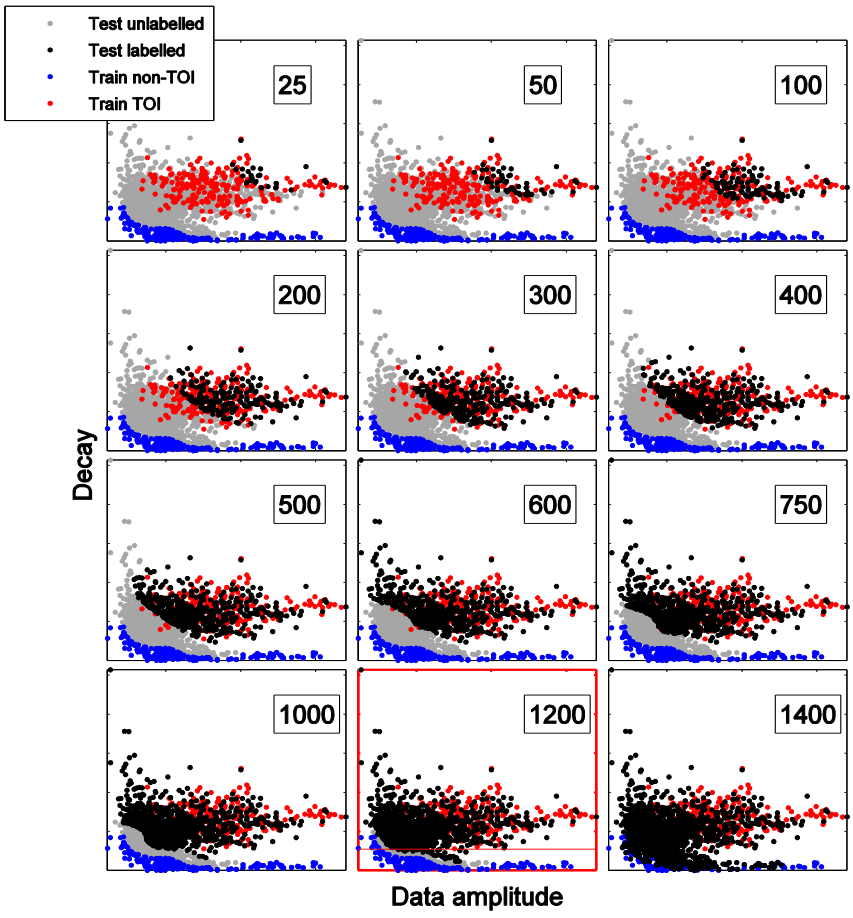


Figure 23. Dig order for PNN classifier applied to Beale EM-61 open field data. The number in the top right of each subplot indicates the first *N* labeled test items found during digging, displayed as

black markers in the test data. Digging in the first stage was terminated after 1200 items (red subplot). Horizontal red line indicates a decay rate cut-off of 0.08.

The training distribution of TOI used feature vectors from the previous ESTCP demonstration at Camp Butner. A similar range of target classes (105 mm, 37 mm, fuzes) was expected at Camp Beale, and there is a good correspondence between the training TOI and large amplitude, slow-decaying feature vectors in the test data. For training non-TOI we use unlabeled test items with the maximum misfit to known TOI as *assumed* non-TOI.

Figure 24 shows a retrospective comparison of EM-61 ROCs obtained with various classification methods for the Beale handheld area. We compare

1. PNN classifier trained on data amplitude and polarizability decay. This was our chosen technique for the Beale EM-61 data sets.
2. PNN classifier trained on polarizability size and polarizability decay.
3. Threshold on polarizability decay.
4. Threshold on data amplitude.

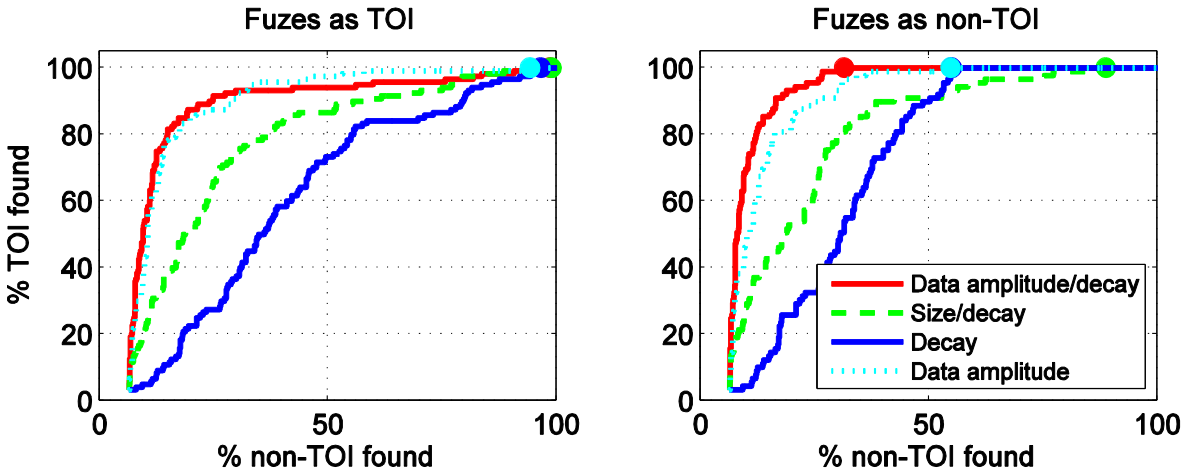


Figure 24. Retrospective ROC curves for classification with EM-61 data, Beale handheld area.

Our usual strategy with EM-61 data is a threshold on decay rate. This method is fairly conservative and the resulting ROCs have relatively poor AUC and mediocre FAR. The PNN classifier trained on data amplitude and polarizability decay achieves the best performance of all methods applied to these data. As might be expected with EM-61 data, all techniques have difficulty identifying small fuzes. The data amplitude/decay classifier does surprisingly well identifying larger TOI, with approximately 30% of clutter excavated in order to find all TOI. This is a marked reduction in the EM-61 false alarm rate for previous demonstrations with comparable TOI classes (e.g. Camp Butner). We attribute this difference to the distributions of clutter: there is a reasonable separation between the distributions of TOI and non-TOI at Camp Beale (Figure 25). At Camp Butner a significant proportion of clutter was of comparable size (and anomaly amplitude) to 37mm items, and so identification of this TOI class was quite challenging. While good classification performance can sometimes be achieved with the EM-61, it is highly dependent on the relative sizes of TOI and non-TOI.

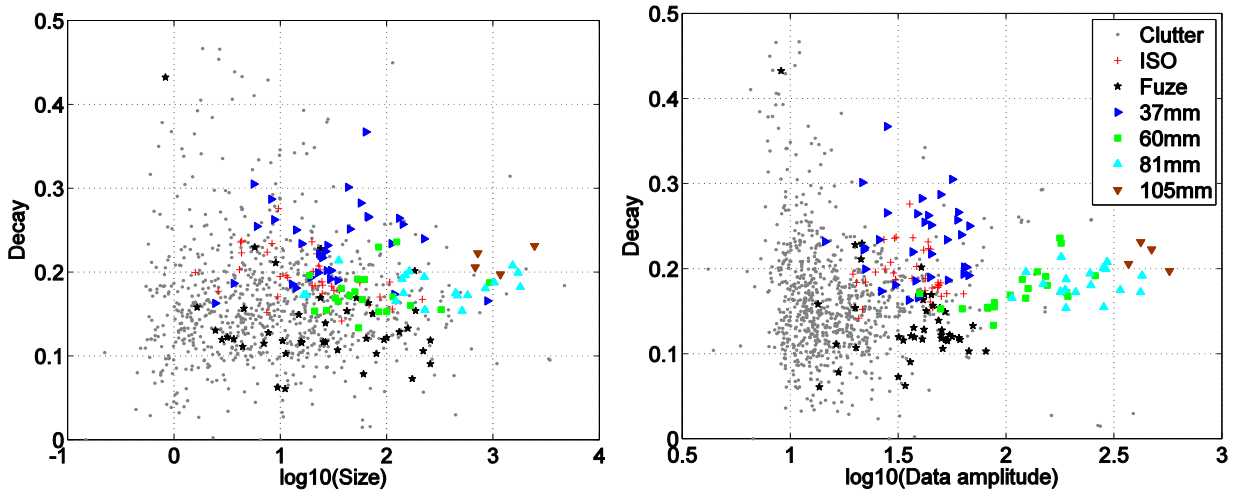


Figure 25. Comparison of polarizability size (left) and data amplitude (right) features for EM-61 data, Beale handheld area.

Another important result from the Camp Beale demonstration is the failure of pre-screening methods used to identify high confidence non-TOI in EM-61 data. We pre-screened a small number (231) of EM-61 anomalies on the basis of their location in data amplitude/decay space. As illustrated in Figure 26, pre-screened test vectors were farthest from TOI training vectors in this two-dimensional space.

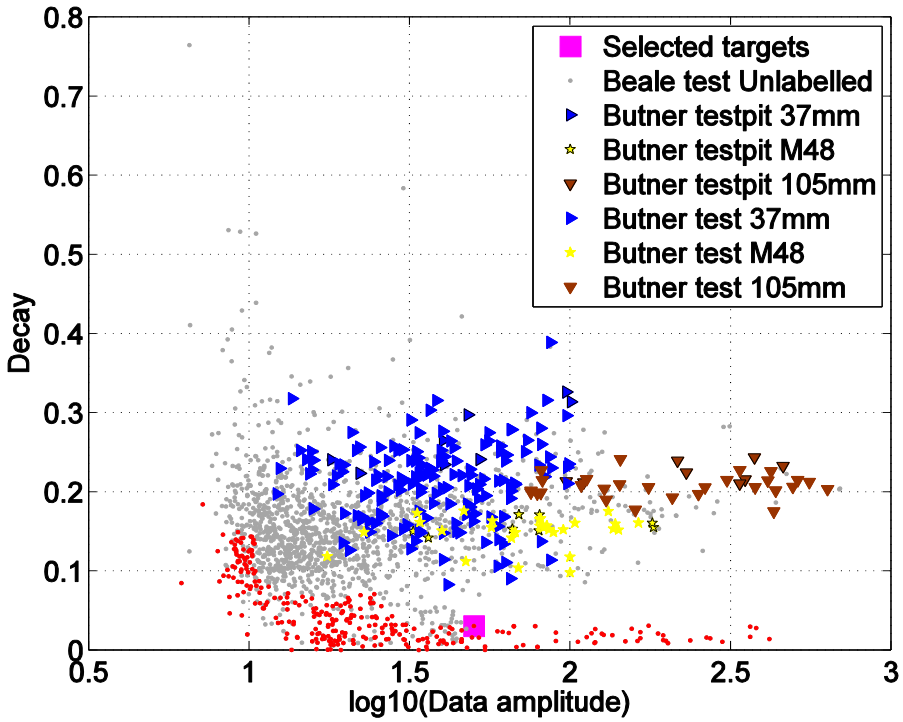


Figure 26. Prescreened feature vectors (red markers) for Beale open area. Target highlighted in magenta is missed seed target 2277.

For our first data request we then assumed that these targets would not be reacquired with the MetalMapper. Despite a very conservative cut-off, the set of pre-screened targets included a seed item (target 2277) and a number of small fuzes. Target 2277 is a 37 mm target with good SNR (~ 30 mV at channel 1), but with an anomalously fast decay rate, as shown in Figure 26. A number of small pieces of barbed wire were also excavated at this location (Figure 27), suggesting that the measured sensor data interrogated these clutter items rather than the target of interest. This is supported by the MetalMapper polarizabilities for target 2277, which look nothing like a 37 mm (Figure 27).

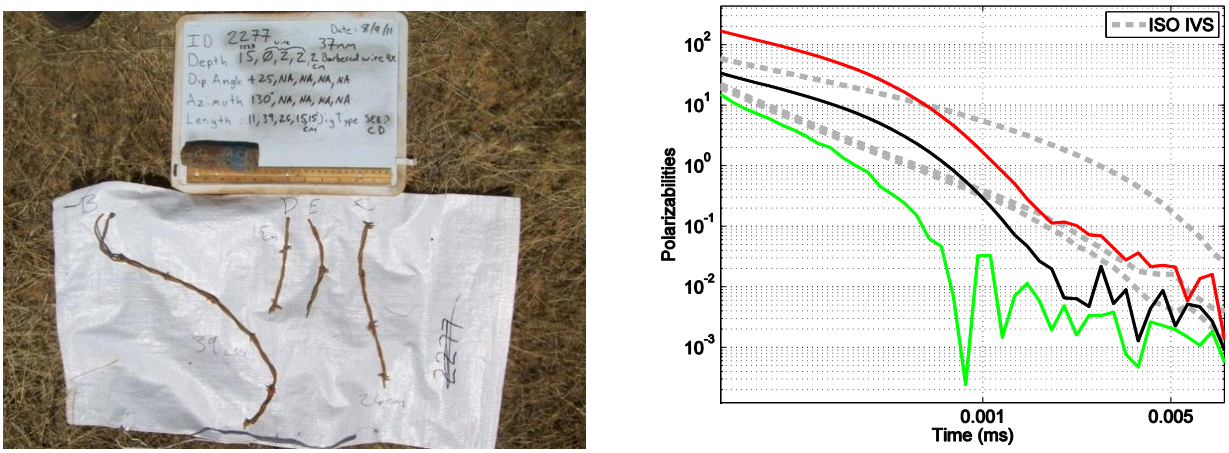


Figure 27. Groundtruth (left) and Parsons MetalMapper polarizabilities (right) for seed item 2277. Target response is likely dominated by wires and recovered polarizabilities do not match with 37mm items.

Even if we account for this possible groundtruth error, native fuzes appear as low SNR, fast-decaying targets in the EM-61 data and cannot be differentiated from fast-decaying clutter. At the 2010 Camp Butner demonstration, pre-screening also mislabeled TOI items. We therefore conclude that pre-screening with the EM-61 is only feasible at sites where targets of interest are large (105 mm, 4.2") and easily distinguished from native clutter. If smaller TOI are encountered, then all detected targets should be reacquired with a cued sensor. In future demonstrations we will not pursue pre-screening techniques.

Appendix B: METALMAPPER FEATURE EXTRACTION AND CLASSIFICATION

There were two MetalMapper datasets acquired over identical anomaly locations by two different contractors at Camp Beale. Two unique Sky analysts processed the respective MetalMapper datasets maintaining a strict firewall throughout the investigation. Results for the two different approaches are presented separately.

Parsons MetalMapper static feature extraction and classification

Background-corrected MetalMapper cued data for 1256 of 1470 anomalies were received as a set of CSV files. Data for the remaining 214 anomalies were not initially requested because they had been pre-screened (i.e., deemed to be high-likelihood non-TOI items) using EM61 data (see Appendix A). Initial feature extraction and classification efforts focused on the pre-screened dataset. After submission of a dig list based on this dataset, with all pre-screened data placed in the "do not dig" category at the end of the dig list, and after learning that the EM61 pre-screening had not been effective, we received the background-corrected data for the 214 pre-screened anomalies. A second round of feature extraction and classification using the full dataset ensued, resulting in submission of a second dig list that did not employ pre-screening.

The data were inverted in UXOLab using a sequential inversion approach to estimate target location, depth and primary polarizabilities. Instrument height above the ground was assumed to be 7 cm. Noise standard deviation estimates were not available, so a constant noise value of 1 over all time channels was used. Target location was constrained to lie between ±0.5 m of the center of the MetalMapper in both X and Y directions. Target depth was constrained to lie between –1.2 and 0 m below the surface. The initial optimization for target location identified up to three starting models to input into the subsequent estimation of polarizabilities. The data for each anomaly were inverted using both a single-object inversion (SI) and multi-object (MI, i.e. two-object) inversion.

Visual QC of the data was performed using a newly developed flexible QC tool which provides a thorough overview of the observed and predicted data, predicted model parameters, and measures of data/model quality. Predicted polarizabilities were compared to reference polarizabilities for various ordnance derived from test pit and IVS measurements. The library of reference items was augmented with additional items based on ground truth obtained through training data requests. During QC the primary objectives were to pass the best model (i.e., the one with the most UXO-like polarizabilities), flag high-likelihood TOI, and fail unlikely models. The latter most commonly applied to one of the MOI models which frequently are unrealistic (e.g., deep, large in magnitude, and sometimes located on or near a horizontal inversion boundary; e.g., Figure 28) but which sometimes provide the best fit to the reference polarizabilities. In many cases the decision on which model to pass was not obvious. In these cases more than one model was passed; the classification procedure would consider all passed models and effectively use the one that is “best”.

The total Beale MetalMapper dataset comprised 1547 anomalies including recollects at some anomalies. The total number of models QCed was 4641 (three per anomaly). Of these, 1602 models were passed and used in the subsequent classification phase. 171 anomalies were identified during QC as potential or very likely TOI.

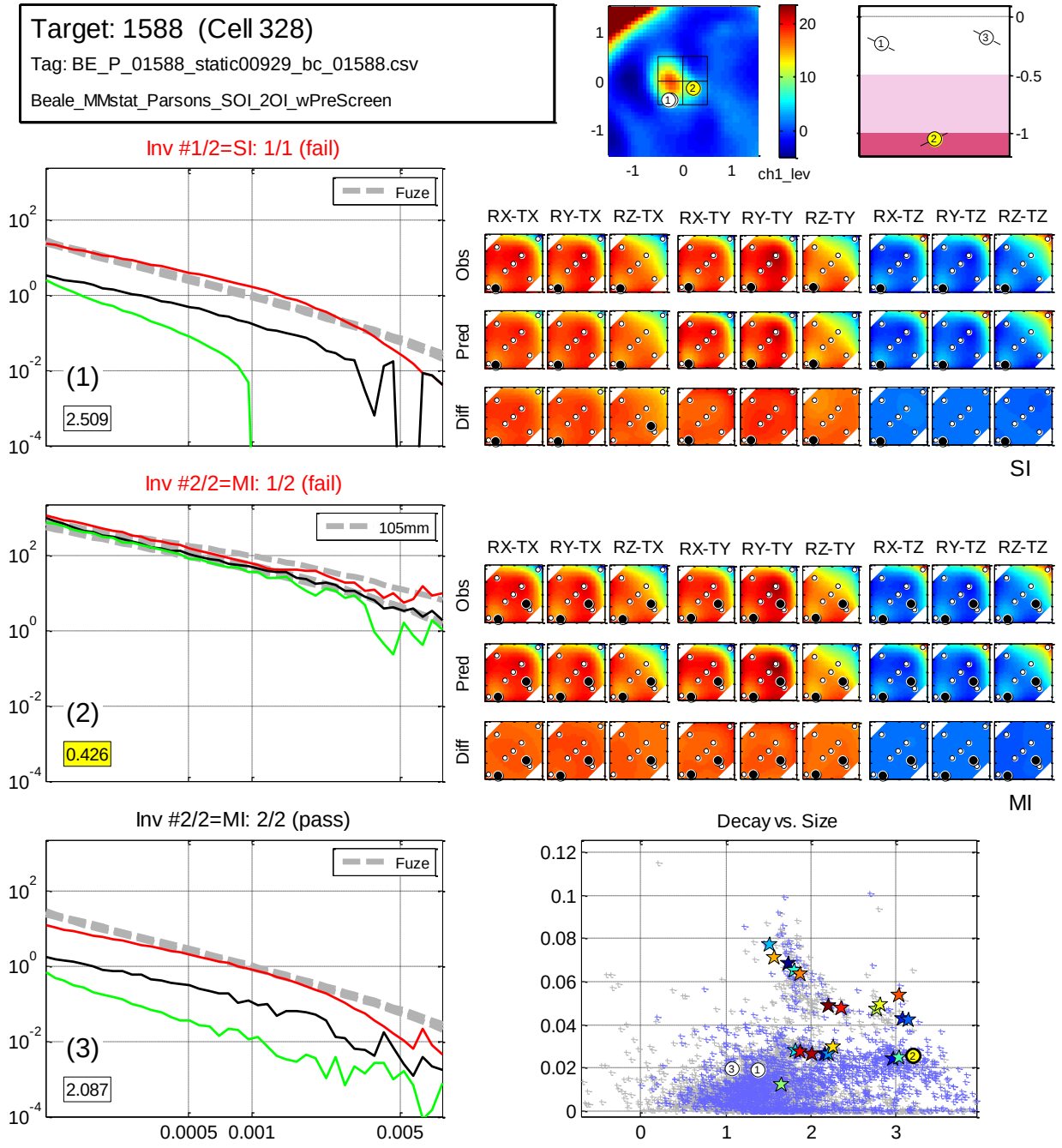


Figure 28. Example of an unrealistic MOI model (anomaly 1588; scrap). The first model of the MOI (model 2) provides (by far) the best fit to the reference polarizabilities (misfit = 0.426), but the predicted depth of 1.05m and high amplitude of the polarizabilities, especially in relation to the relatively weak EM61 anomaly, are indicators that this model is an artifact of the MOI process. Accordingly, this model was failed during QC.

Training data selection

Figure 29 shows the distribution of all passed models in decay-size feature space.

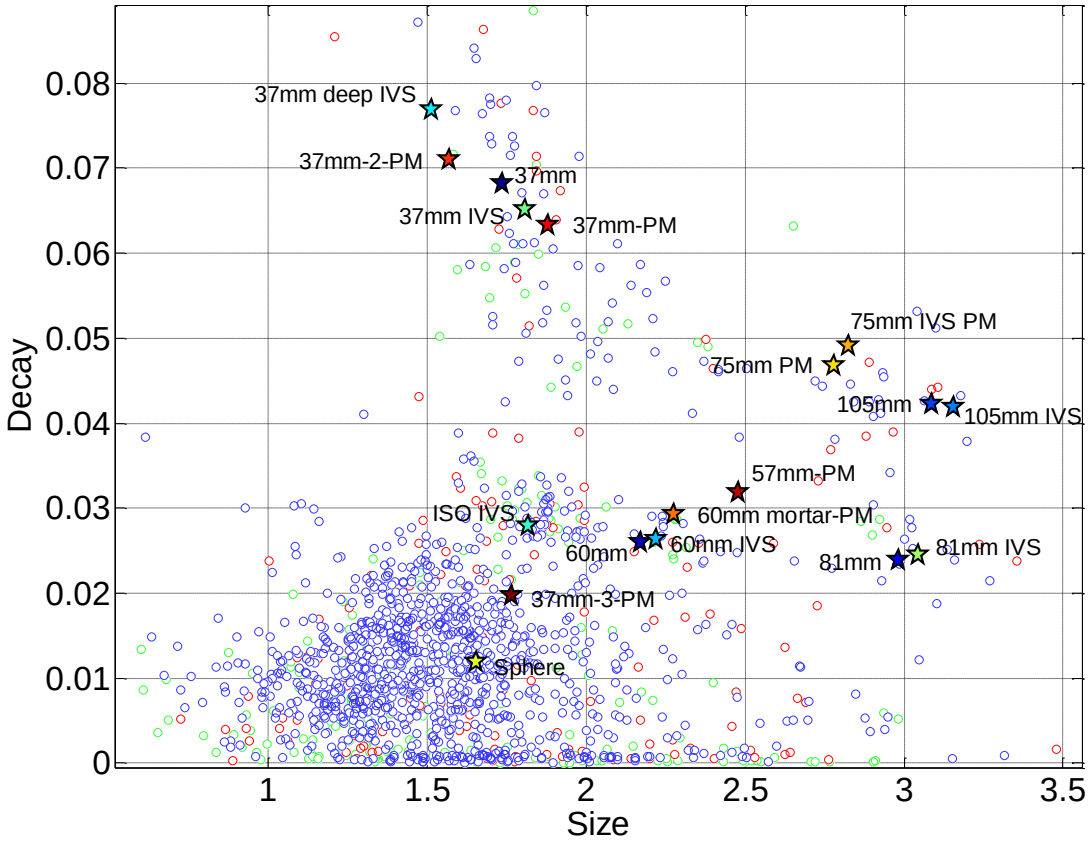


Figure 29. Distribution of all passed models in decay(t1,t29) versus size(t1) feature space, where size(t1) is the total polarizability measured at the first time channel (t1=0.106ms), and decay(t1,t29) is size(t1)/size(t29) where t29=2.006ms. Circles, squares and triangles correspond to SOI and MOI (models 1 and 2), respectively. Labeled stars show location of library reference items in feature space.

A newly developed training data selection tool was used to assist with the selection of training data. This tool is used to find clusters of items with similar polarizabilities. A basic example of its use is shown in Figure 30 and Figure 31. When dealing with clusters of items with polarizabilities similar to those of a known reference item, the general approach was to request training data for items with polarizabilities that were not too similar to those of the reference item (e.g., items with noisy secondary polarizabilities, or with some separation between the secondary and tertiary polarizabilities). This approach provides information on the typical range of variation in polarizabilities of TOI, thus providing a sense of the ability of the data to discriminate TOI from non-TOI.

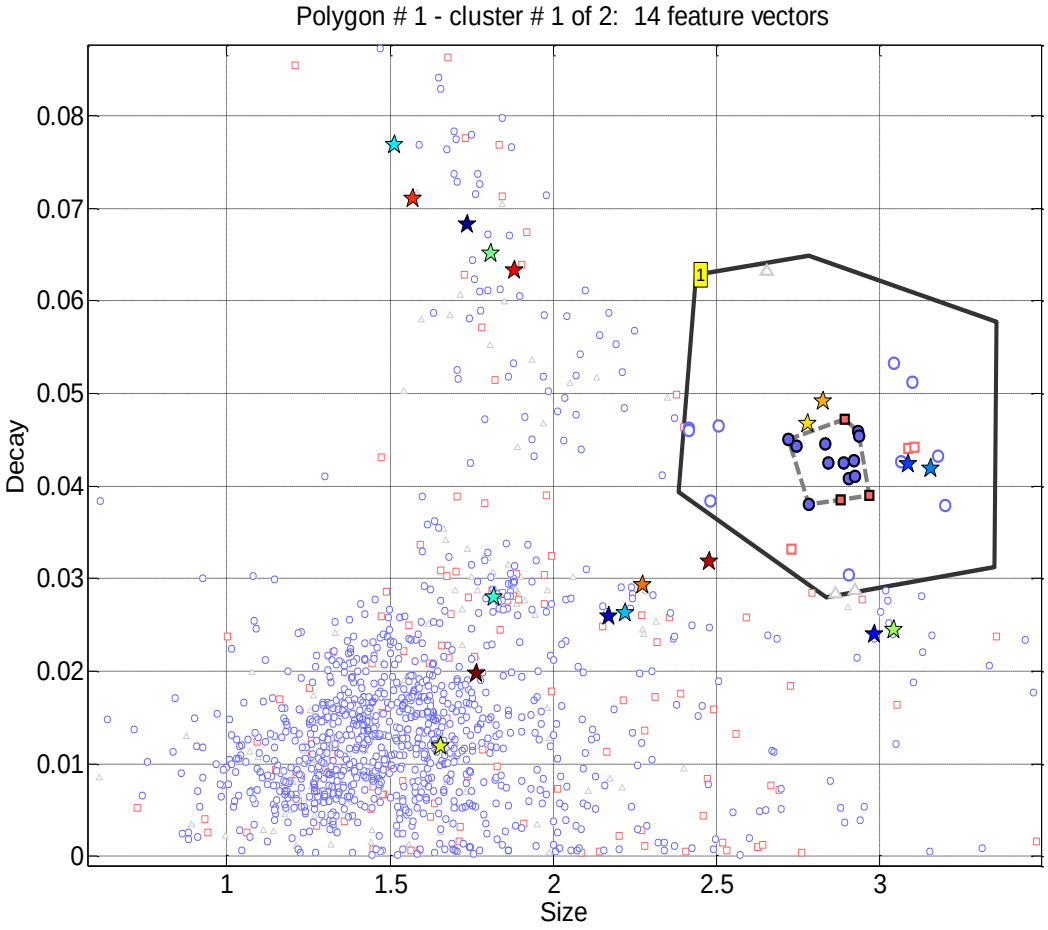


Figure 30. Example of use of training data selection tool. A polygon (heavy black line) is drawn in feature space. Clusters of items with self-similar polarizabilities are automatically found. In this case two clusters were found within the polygon; one is visible (solid feature symbols surrounded by broken grey line). Polarizabilities for this cluster are shown in Figure 31.

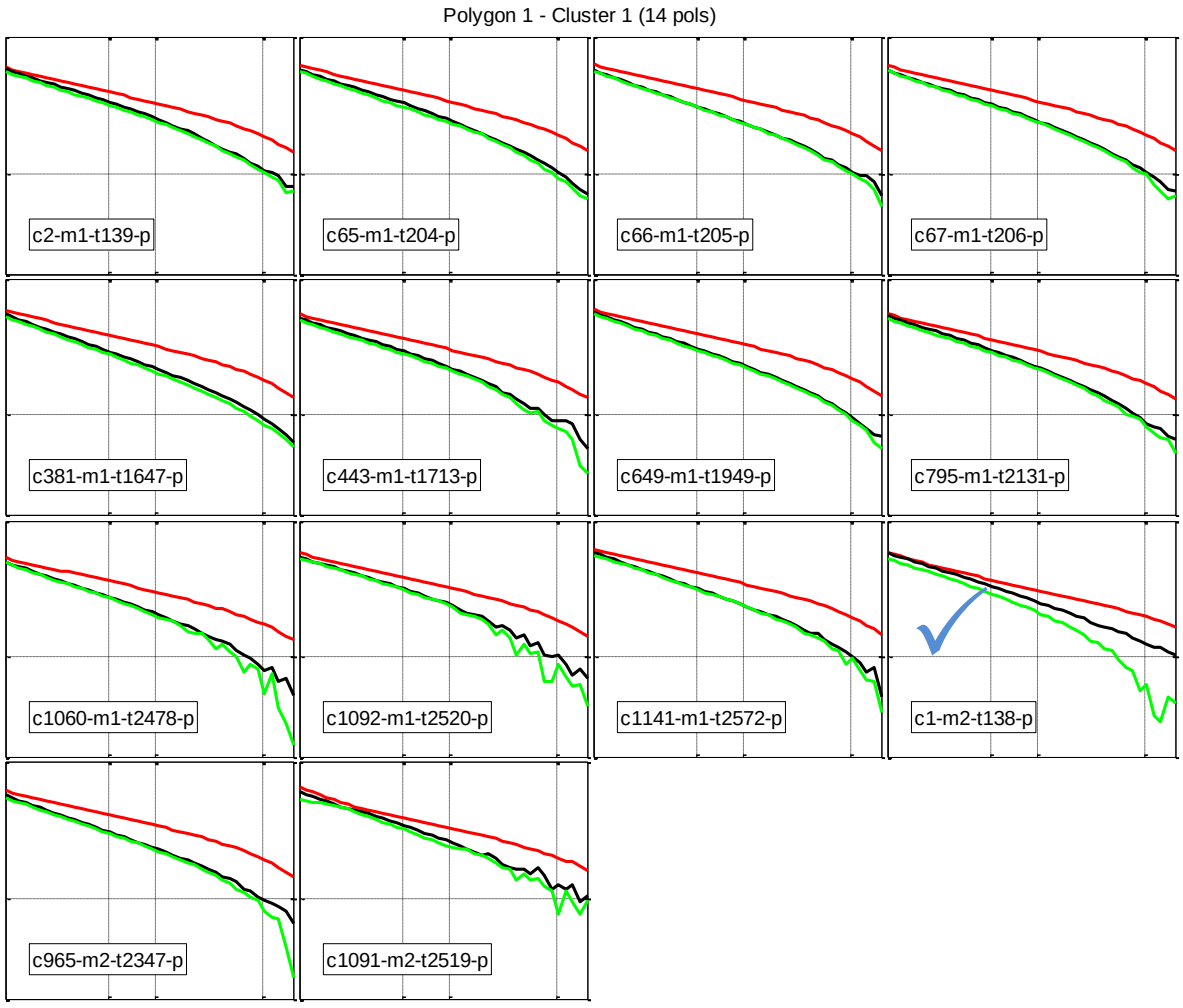


Figure 31. Polarizabilities for the cluster shown in Figure 30. Training data were requested for Anomaly 138 (check marked); ground truth showed that this anomaly corresponds to a 105mm TOI.

In some cases training data were not requested for a cluster of items. E.g., Figure 32 and Figure 33 show a cluster of items which, based on previous experience from Pole Mountain, all likely represent horseshoes.

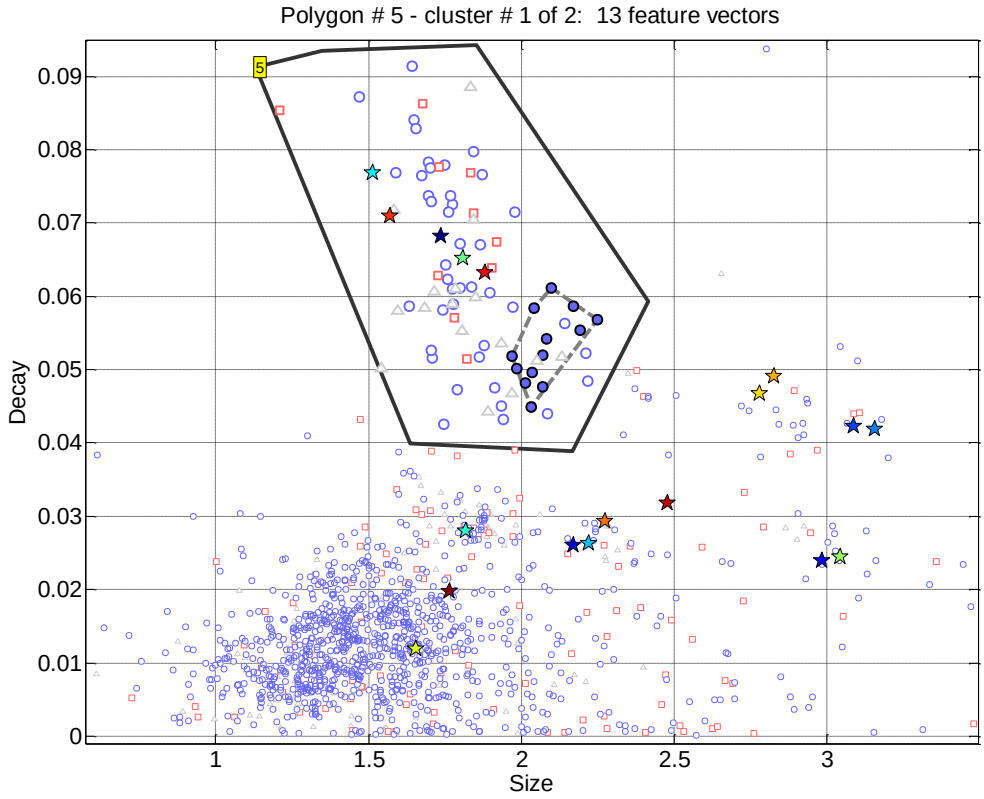


Figure 32. Cluster with polarizabilities different from those of all reference items. Items in the cluster are solid feature symbols surrounded by broken grey line. Polarizabilities for this cluster are shown in Figure 33.

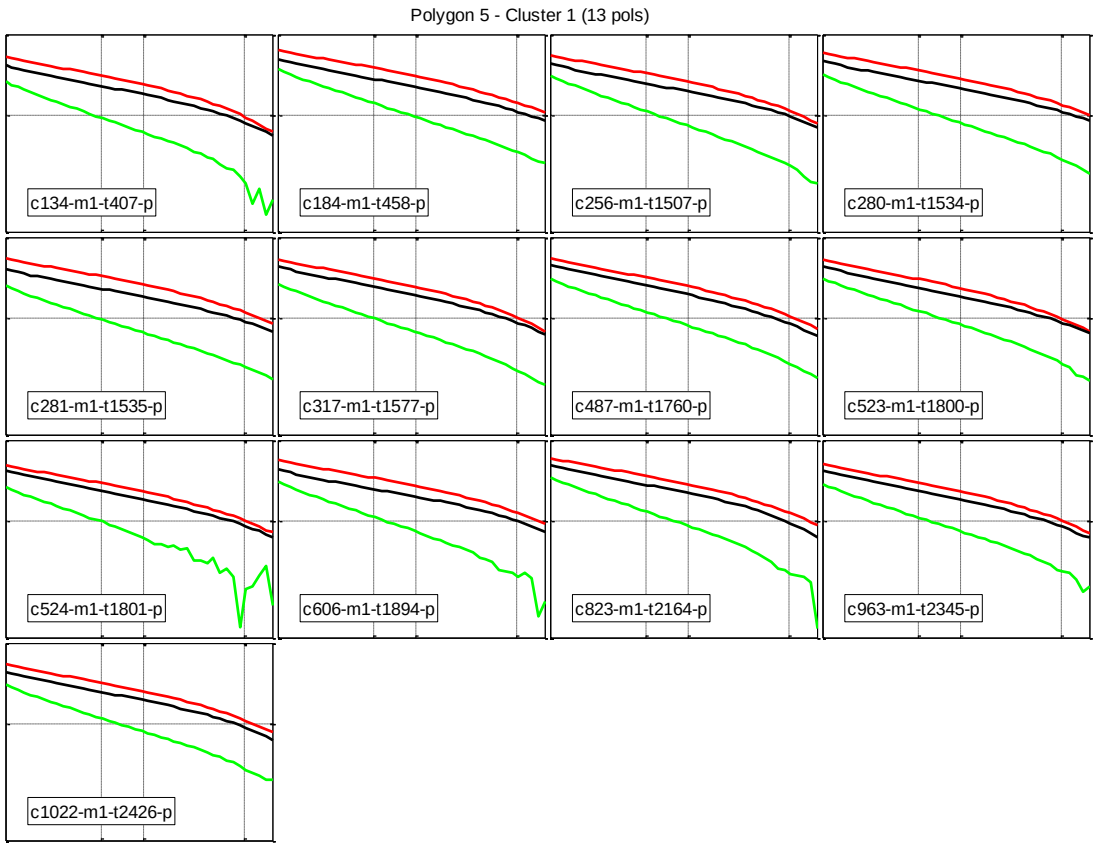


Figure 33. Polarizabilities for the cluster shown in Figure 32. The polarizabilities are not UXO-like and based on past experience were recognized as representative of horseshoes. Training data were not requested for any of these items.

In addition to clusters corresponding to known items (e.g., 37mm, 60mm, 81mm, 105mm and small ISO), a fairly large cluster of items with sphere-like polarizabilities was found (Figure 34and Figure 35).

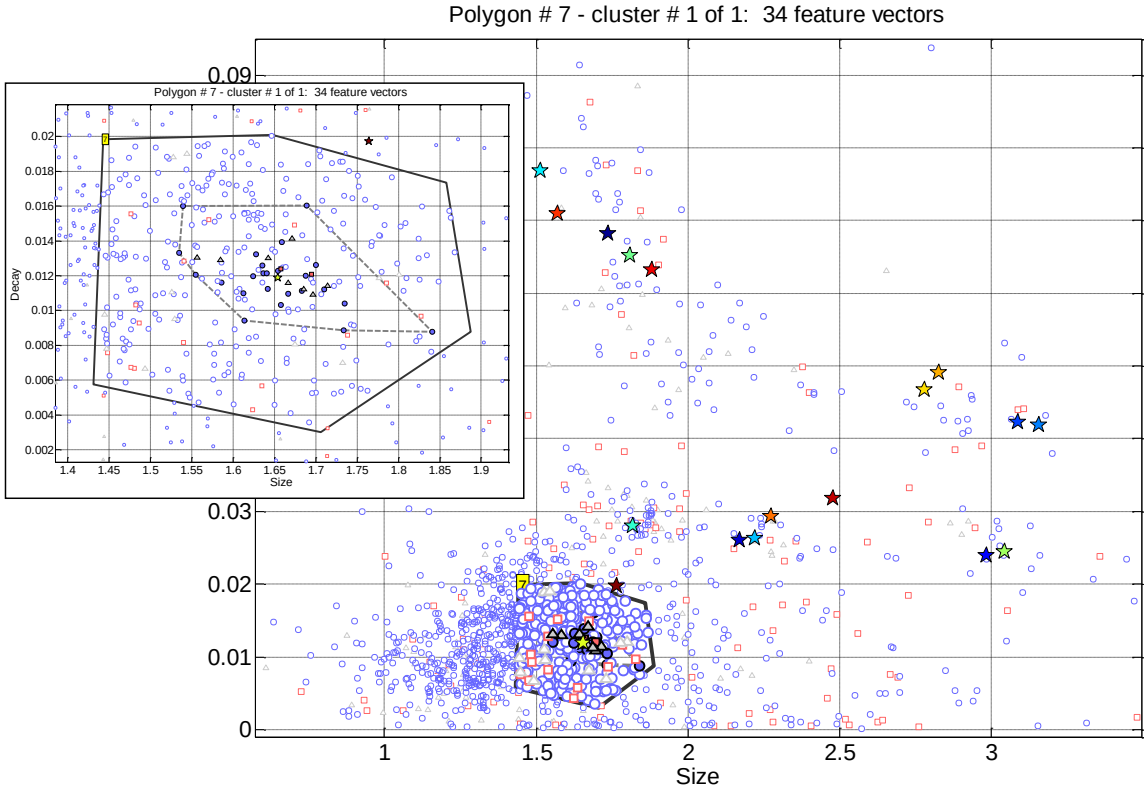


Figure 34. Cluster with polarizabilities similar to those of a sphere. Items in the cluster are solid feature symbols surrounded by broken grey line. Inset shows close up of the polygon surrounding the cluster. Polarizabilities for this cluster are shown in Figure 35.

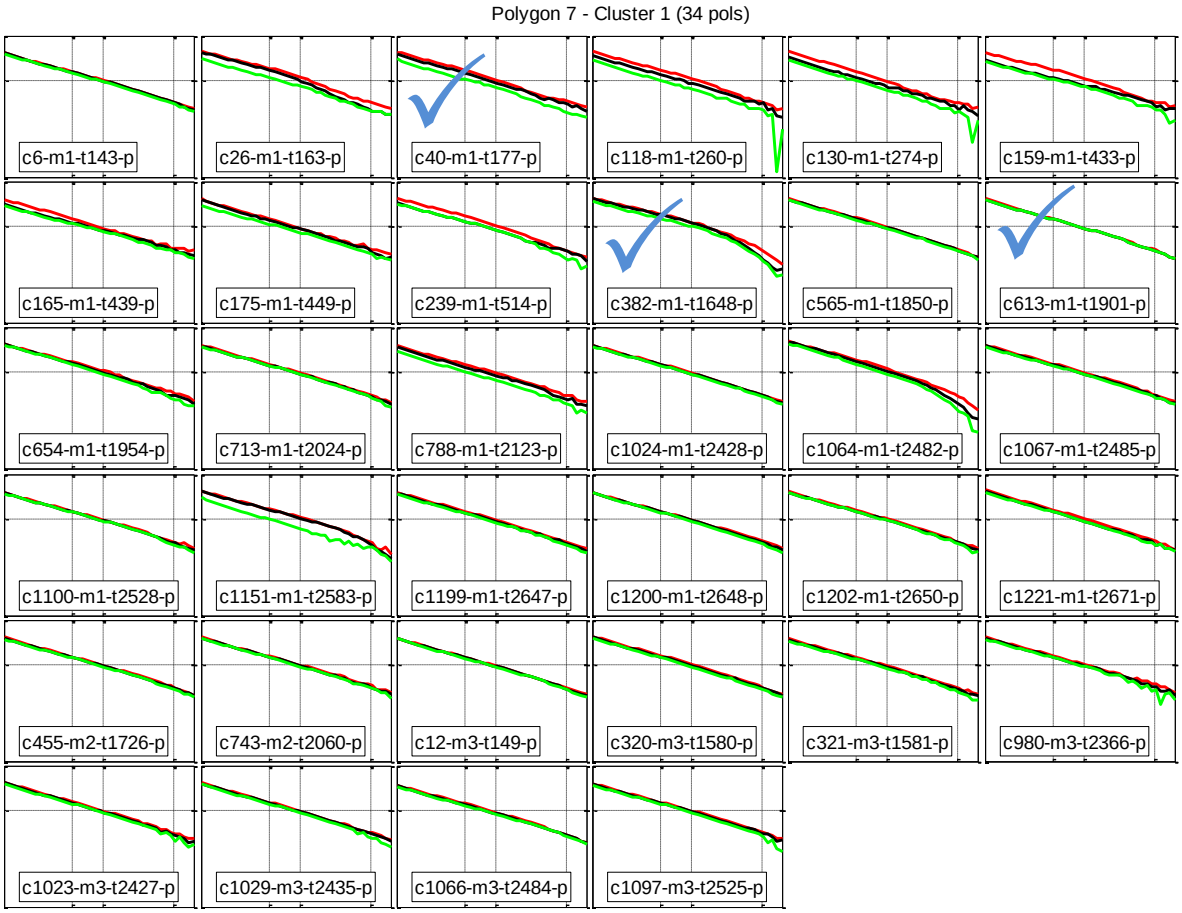


Figure 35: Polarizabilities for the cluster shown in Figure 34. Training data was requested for the three checked items: anomaly numbers 177, 1648 and 1901. Two of these (177 and 1648) turned out to be scrap. Anomaly 1901 is a fuze part, initially considered by the program office to be a TOI.

Training data were requested for 14 items (Table 8). Of these, nine were TOI.

Table 8. Training data requests. Items highlighted in yellow are TOI.

Anomaly	Depth	Dip Angle	Azimuth	Identification	Length	Dig Type
BE-138	36	36	52	105mm	65	TOI
BE-141	25	35	182	60mm	13	TOI
BE-1759	30	-30	60	60mm	24	TOI
BE-2643	1	0	9	Scrap Metal	15	CD
BE-1807	18	-20	325	ISO Item	10	TOI
BE-2645	13	-90	n/a	37mm TP-T	11	TOI
BE-1768	7	-8	310	Frag	5	MD
BE-2354	3	9	22	Frag	17	MD
BE-2575	21	60	240	ISO Item	10	TOI
BE-142	13	49	269	ISO Item	10	TOI
BE-1648	2	90	n/a	Tailboom	8	MD
BE-177	7	-37	324	Frag	6	MD
BE-1901	1	n/a	n/a	Fuze Part	6	TOI
BE-167	20	18	295	37mm	11	TOI

Classification method

A new dig list tool (Figure 36) was used to determine the digging order. For each anomaly a score calculated based on a combination of parameters:

- S_{L23} = a measure of the size of the secondary and tertiary polarizabilities L2 and L3 (S_{23});
- M = minimum misfit with all library reference polarizabilities;
- D = a measure of the decay of the total polarizability;
- S_{Ltot} is the size of the total polarizability;
- $\overline{M}$ = minimum misfit with a set of "non-ordnance" reference polarizabilities (e.g., horseshoe);
- Q = polarizability quality. Polarizability quality is an *ad hoc* measure of (1) how much the polarizabilities looks like those of a typical UXO (i.e., polarizabilities of an axi-symmetric body); and (2) the smoothness of the polarizabilities.

This is the approach that was taken with the Beale Parsons MetalMapper data.

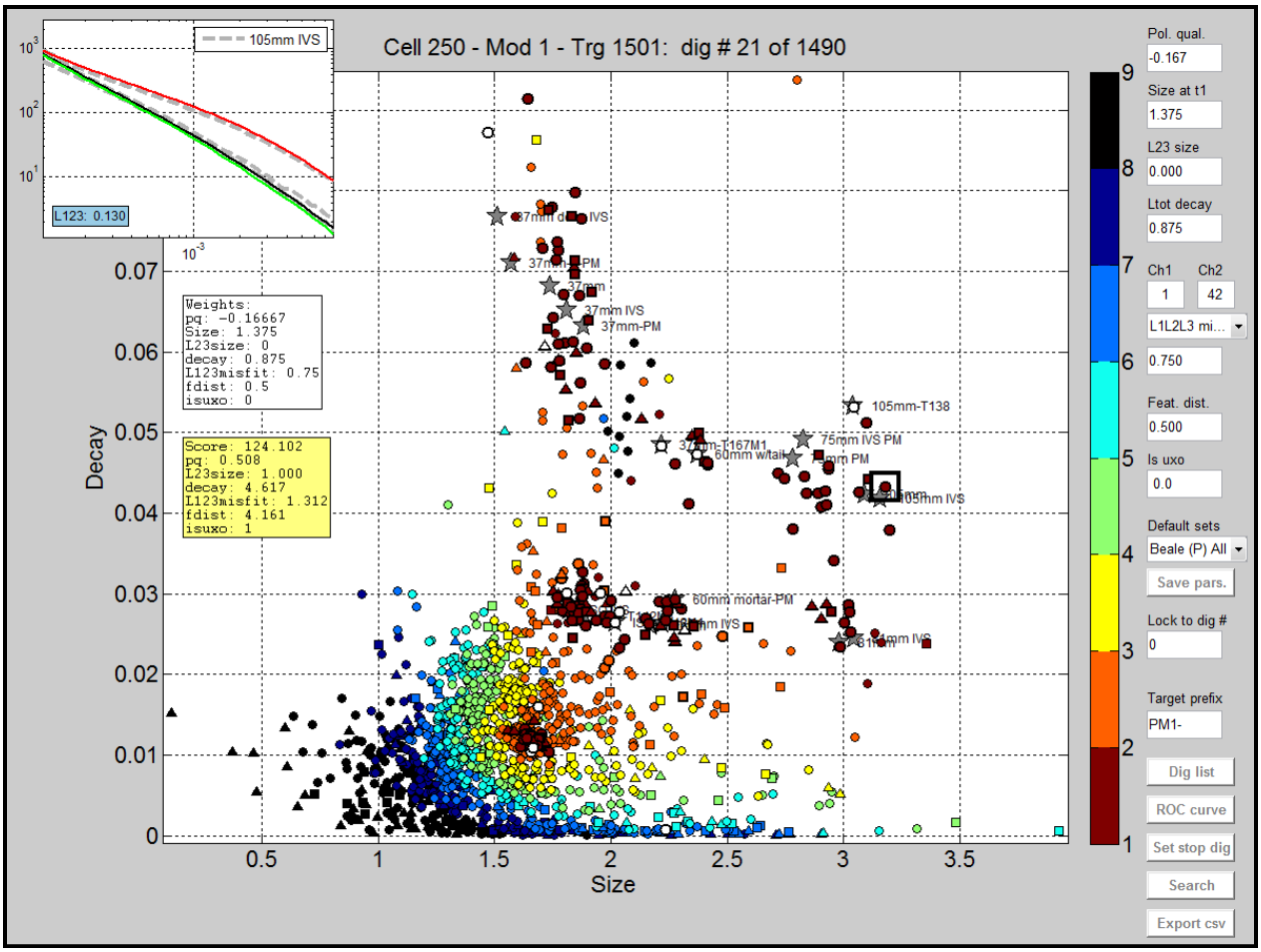


Figure 36. Dig list tool graphical user interface. Features are plotted in decay-size space with each feature color-coded according to its location in the dig list (red earliest; black latest). Features colored white are training items. Inset at top left shows polarizabilities of the selected anomaly (feature surrounded by a black square). The dig list order is based on the weights shown on the right. These can be specified manually or optimal weights can be determined by a search procedure. The latter approach was used for the Beale Parsons MetalMapper data.

A stop dig point was determined by visual inspection of the predicted polarizabilities (in relation to the best fitting reference polarizabilities) of each anomaly plotted in dig list order (Figure 37). The stop dig point was conservatively set to the latest anomaly in the dig list with polarizabilities judged to have a realistic possibility of corresponding to a TOI.

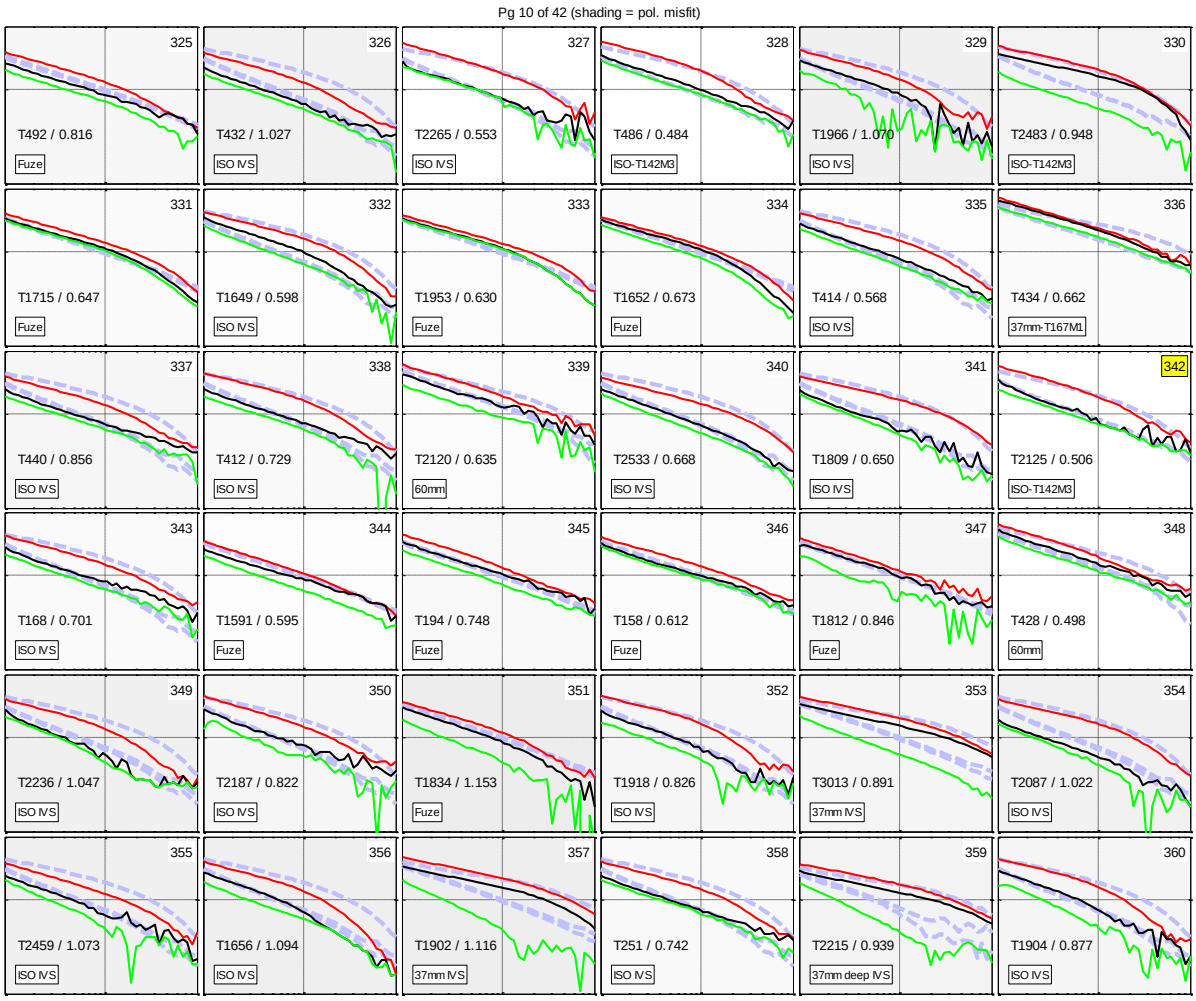


Figure 37. Example display of polarizabilities plotted in dig list order. Here, polarizabilities for dig numbers 325 through 360 are shown. Number in top right of each panel is dig number. Numbers in lower left are anomaly number (preceded by "T") and misfit (calculated using all three polarizabilities) to best fitting reference polarizabilities (broken light blue lines). Text in lower left corner is name of closest fitting reference item. Each panel is shaded according to misfit, with larger misfits corresponding to darker shading. Yellow-highlighted dig number (342) is the stop dig point that was chosen for the dig list in which pre-screening was not employed.

Figure 38 shows the partial ROC curve obtained from the program office for the dig list with no pre-screening. At the time of the scoring, fuzes were considered TOI. This result looked excellent: the partial ground truth showed that the last TOI found was at dig 212, followed by 130 non-TOI digs. The number of digs necessary to achieve a specified confidence level that all TOI have been found can be estimated by fitting a bi-normal distribution to the observed partial ROC curve (Figure 39). At the 99% confidence level we found no further digs were required. Based on this, no modifications were made to the originally submitted dig list.

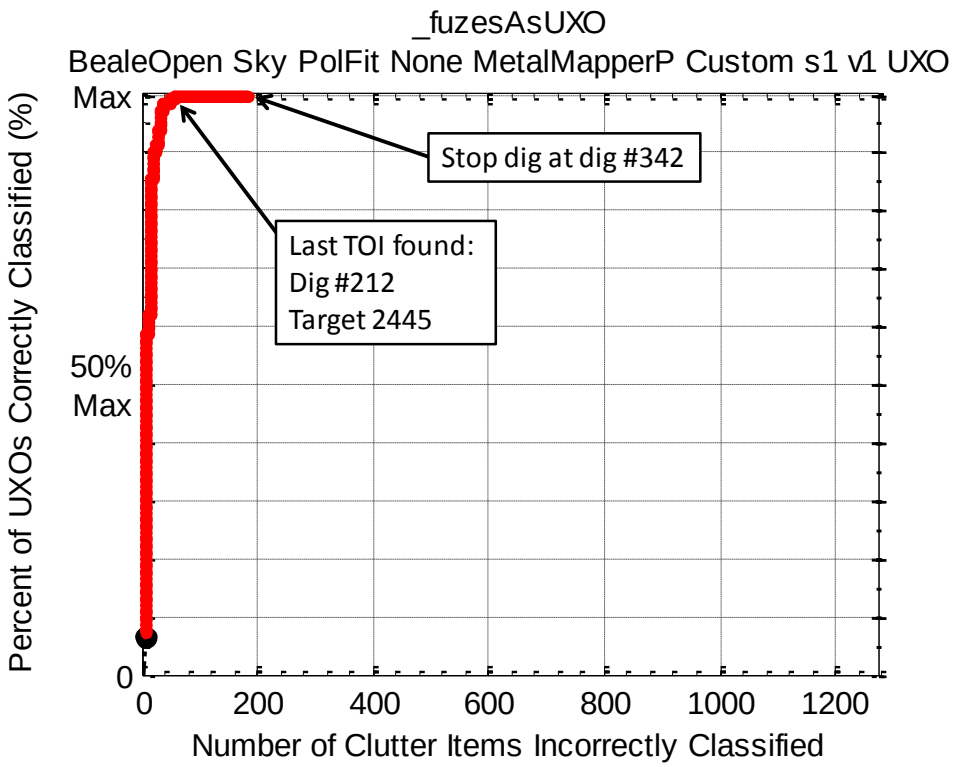


Figure 38. Partial ROC curve for the dig list with no pre-screening. Fuzes are considered TOI.

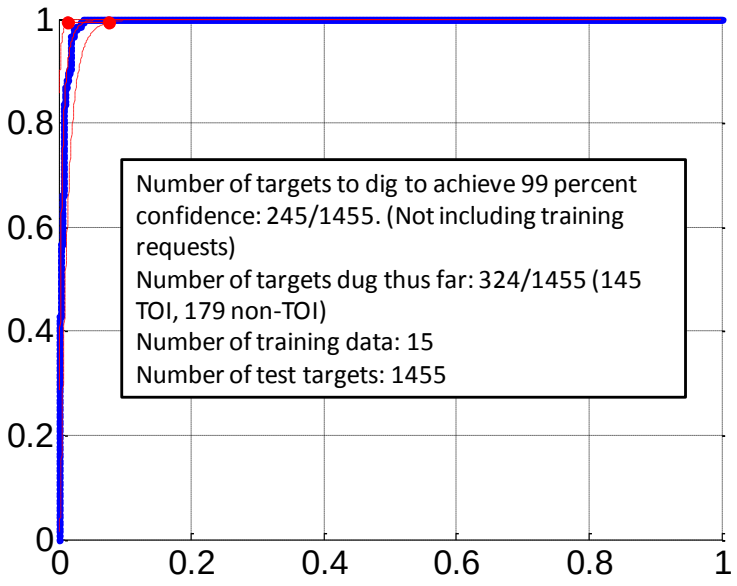


Figure 39. By fitting a bi-normal distribution to the observed partial ROC curve the number of digs necessary to achieve a specified confidence level (that all TOI have been dug) can be estimated. At the 99% confidence level, our partial results suggested that no further digging was necessary.

Parsons MetalMapper static retrospective analysis

Because pre-screening using EM61 data did not prove effective (e.g., a 37mm item - anomaly 2277 - was missed), we limit our retrospective analysis to the dig list which did not employ EM61 pre-screening. Figure 40 shows the location in decay versus size feature space of all TOI for the Parsons MetalMapper dataset.

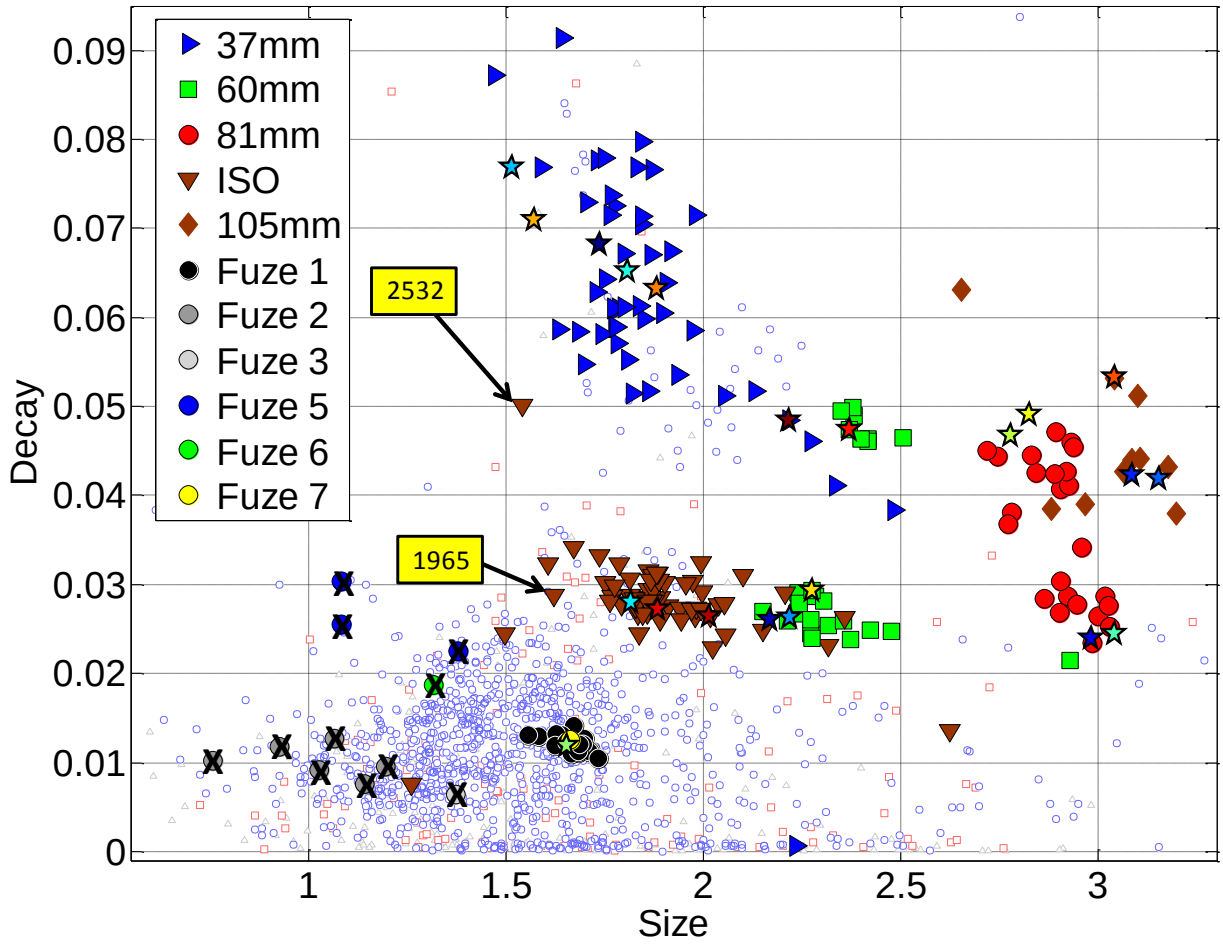


Figure 40. Decay versus size feature plot for Parsons MetalMapper data. Large symbols show locations of TOI in feature space. Two ISOs that were missed (anomalies 1965 and 2532) are identified. Eleven small fuzes marked with "X" were missed (for the case of fuzes treated as TOI).

Figure 41 shows the final ROC curve for the case of fuzes considered as TOI. Two ISOs and eleven small fuzes were missed. In the Beale open area there were a total of six fuze types. During training data selection only the large (6cm) fuze was discovered (see Figure 43). Because of this, and because the polarizabilities of the missed fuzes were significantly different from the polarizabilities of all items in the reference library, the classification procedure employed had effectively no chance of finding the small fuzes. This highlights the importance of careful selection of training data, particularly if very small items

are to be found. We requested training data for fifteen items. Significantly more training data would have been required to maximize the likelihood of finding all the small fuzes. Figure 42 and Figure 43 show the polarizabilities and ground truth photos for the missed fuzes.

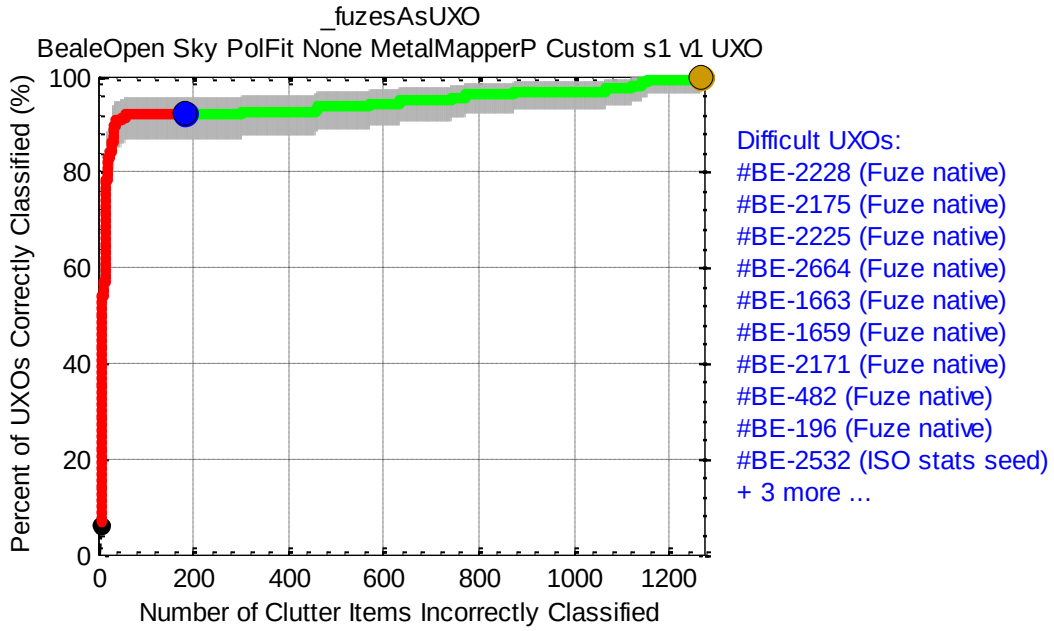


Figure 41. Final ROC curve for the case of fuzes treated as TOI.

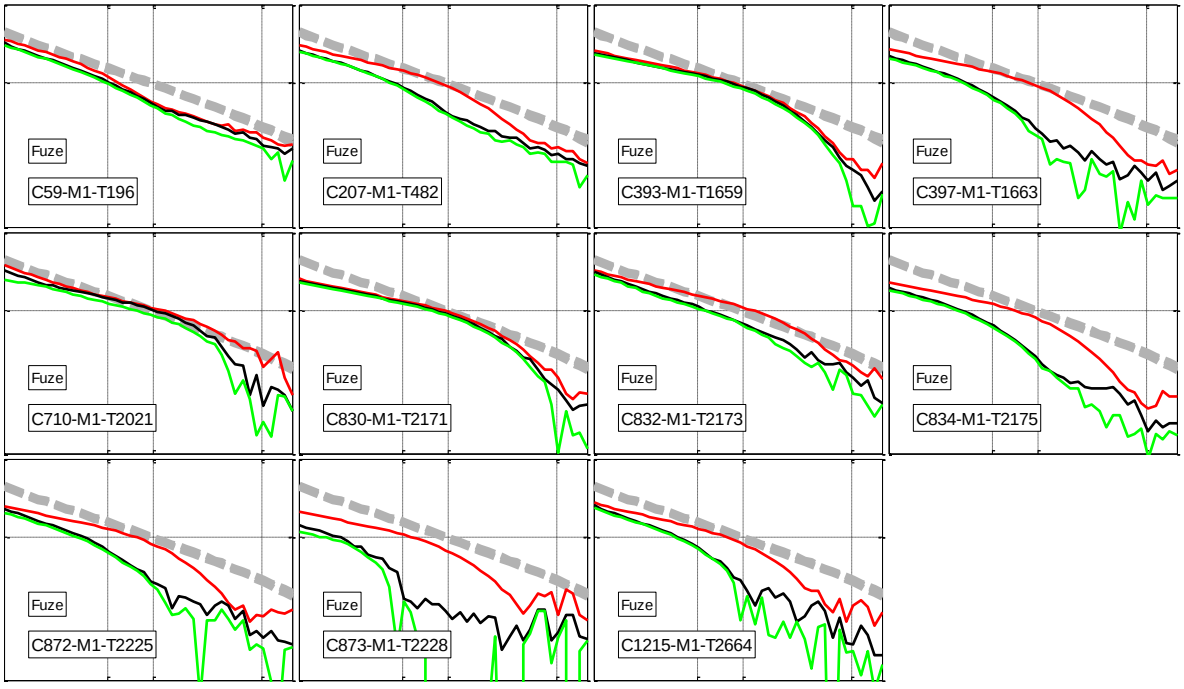


Figure 42. Polarizabilities of the eleven fuzes that were missed for the case of fuzes considered as TOI.

Figure 43. Ground truth photos of the eleven fuzes that were missed for the case of fuzes considered as TOI.

Although all dig lists were scored with fuzes treated as TOI, the program office decided in the end to treat fuzes as clutter. Figure 44 shows the final ROC curve for this case. Two small ISOs were missed (anomalies 1965 and 2532; Figure 45). It is clear from the polarizabilities shown in Figure 45 that had a strategy of switching to matching on the primary polarizability been employed the dig list would have been more effective at digging all TOI early. To test this, we constructed a dig list (treating fuzes as clutter) following the same procedure used to construct the final submitted dig list shown in Figure 44. The first 211 digs were then locked and subsequent dig order was based strictly on the misfit between the observed and best-fitting reference primary polarizability. The resulting ROC curve (Figure 46) shows that all TOI would be found after 262 digs (126 non-TOI digs). Perhaps more surprisingly, a dig list constructed solely on the primary polarizability match (Figure 47) does slightly better, with all TOI found after 260 digs (124 non-TOI digs). This suggests that, while in a few cases the secondary polarizabilities of the Parsons MetalMapper data are unreliable for classification, the data were of sufficient quality to provide good constraints on the primary polarizability of all TOI. This simple approach would not be totally successful with, for example, the CH2M Hill MetalMapper dataset.

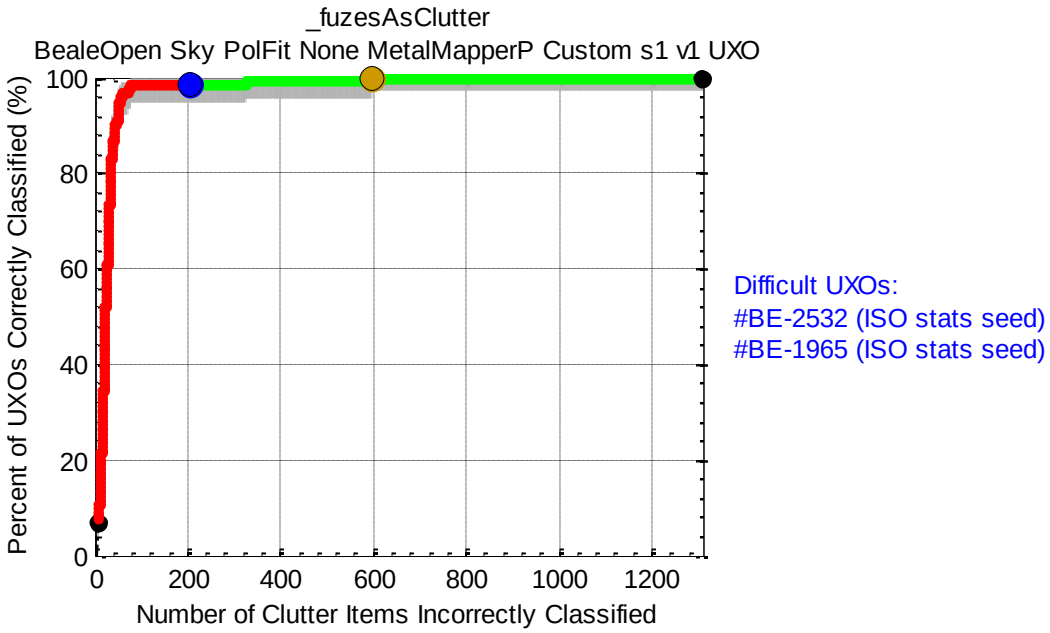


Figure 44. Final ROC curve for the case of fuzes treated as clutter.

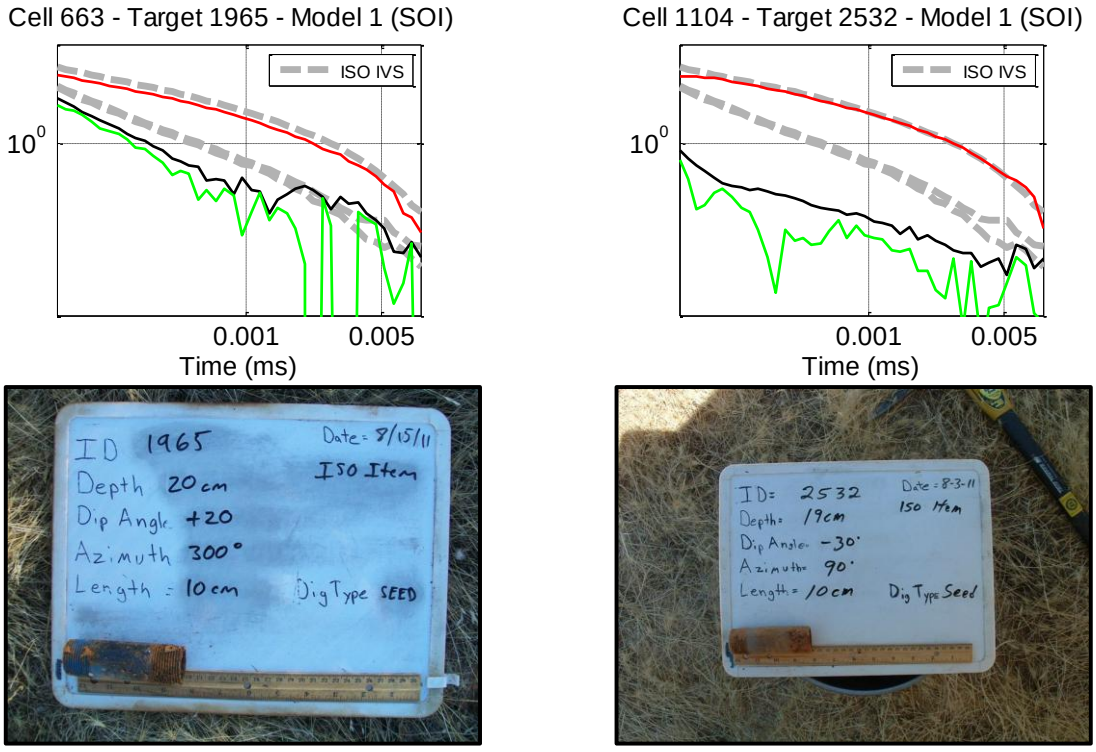


Figure 45. Polarizabilities and ground truth photos of the two ISOs that were missed. Anomaly 1965 appeared at dig 447; anomaly 2532 appeared at dig 728.

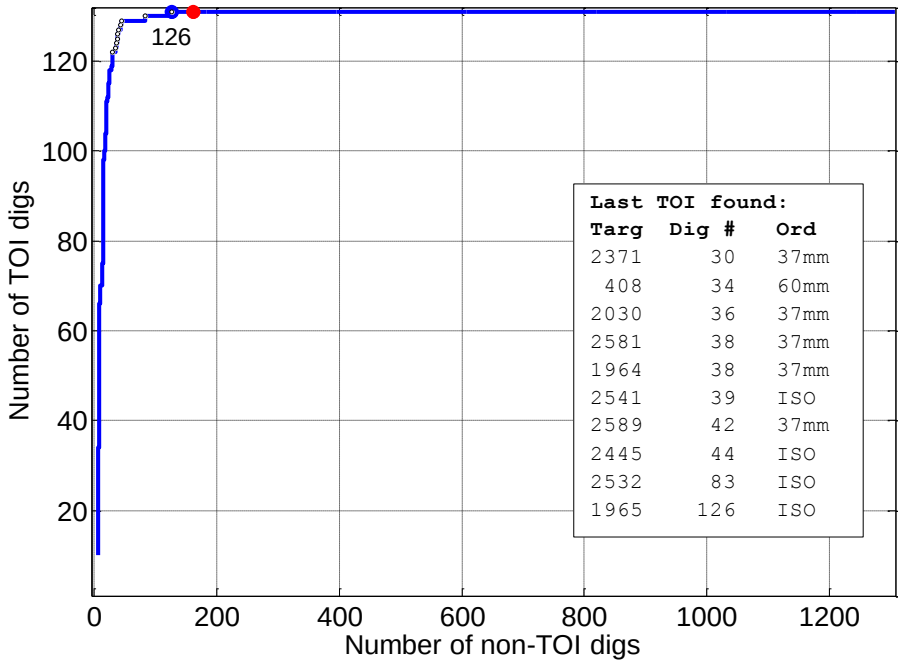


Figure 46. Retrospective dig list (fuzes considered as clutter) constructed following the same procedure used to construct to the dig list shown in Figure 44, but changing to matching on the primary polarizability after 211 digs. All TOI are dug after 262 digs (126 non-TOI digs).

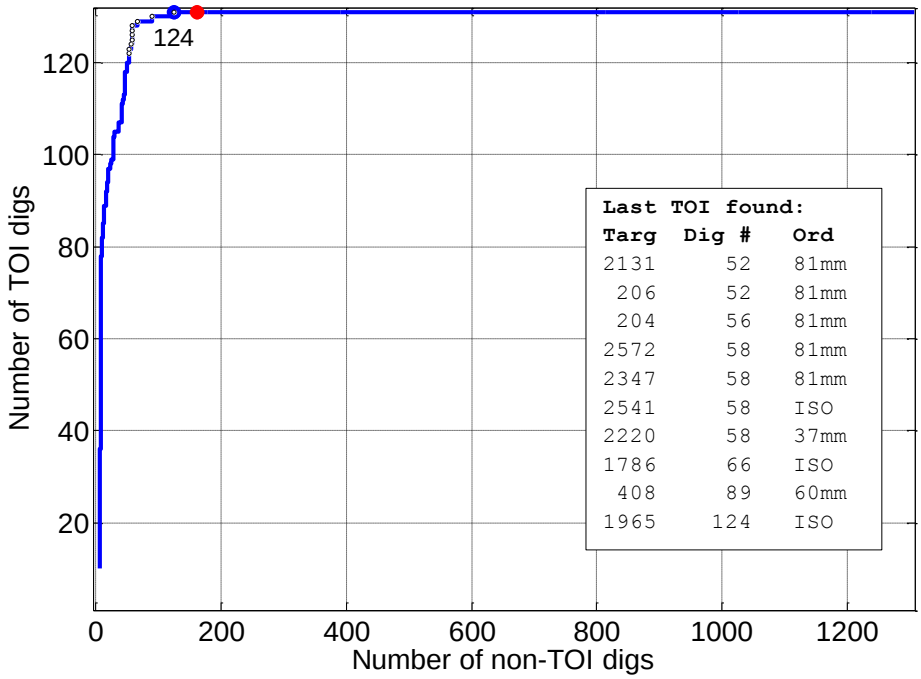


Figure 47. Retrospective dig list (fuzes considered as clutter) based strictly on match to the primary polarizability. All TOI are dug after 260 digs (124 non-TOI digs).

CH2M Hill MetalMapper Static Feature Extraction and Classification

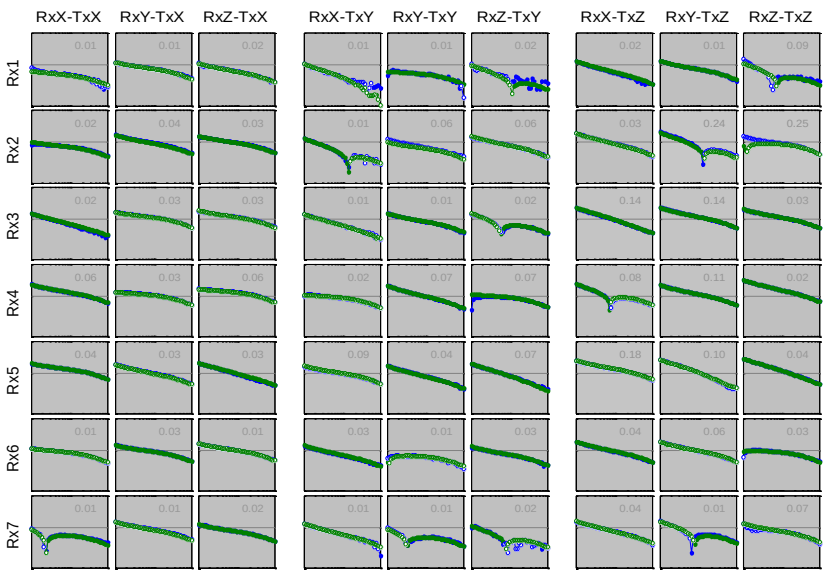
Raw dynamic MetalMapper data with background response removed were received as a collection of CSV files for each of the cued targets. The data were imported and inverted in UXOLab using a sequential inversion approach to estimate target location, depth and primary polarizabilities. Instrument height above the ground was set to 7cm. Noise standard deviations were estimated from test-strip measurements. Target location was constrained to lie between +/- 0.5m in both X and Y directions. Target depth was constrained to lie between -1.2 and 0 m. The initial optimization for target location identified up to three starting models to input into the subsequent estimation of polarizabilities. For each cell of data we performed both a single-object inversion (SI) and multi-object (MI, i.e. two-object) inversion.

A new QC tool was developed allowing inspection of all components of the observed and predicted data and residuals. Examples of data fits and results are shown in Figure 48 to Figure 50. Information available through the QC tool include (1) predicted polarizabilities compared to library polarizabilities for ordnance derived from IVS and test pit data; and (2) a feature space plot showing the predicted location in feature space relative to the predicted locations for all other targets in the data set and the library items. The location of the MetalMapper sensor is also overlain on a gridded image of the EM61 data to gauge how well the sensor is positioned over the anomaly picked from the EM61 data.

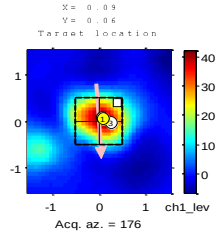
The inverted MetalMapper data set was QCed by a single Sky analyst. Primary QC objectives for each cell were:

1. Pass or fail each inversion (SOI and MOI). All SOI and MOI Inversions were passed for this analysis.
2. For each passed inversion, specifically pass or fail for each model. For example, if one of the MOI models would clearly not be associated with the picked target; that model would be failed.
3. Flag passed inversions as likely UXO (based primarily on visual resemblance of observed and library polarizabilities and location in feature space).

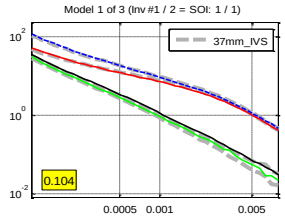
Figure 48 shows results for a high amplitude anomaly readily identified as a TOI (target 2529, 37mm projectile). Figure 49 shows results for a TOI (target 2136, 60mm mortar) here the SOI is not the best fitting model. Primary and secondary polarizabilities for the SOI resemble the reference 81mm at early times but diverge from the reference at late times. The third polarizability is not well recovered for the SI. The predicted locations for the SOI and the second model of the MOI both place the target near the corner of the MetalMapper sensor, offset from the target visible in the EM61 data. These predicted locations correspond to another target in the vicinity of the sensor and those models are therefore failed. Figure 50 shows that the first model of the MOI predicts polarizabilities that match well with the reference 60mm mortar and are recovered at a target location that coincides with the anomaly picked from the EM61 data. This MOI model is passed and will be the only model used to classify target 2629. In other cases where the correct model choice may not be as obvious, both the SOI and MOI models were passed and used in the classification routines.



(a) Data Fit

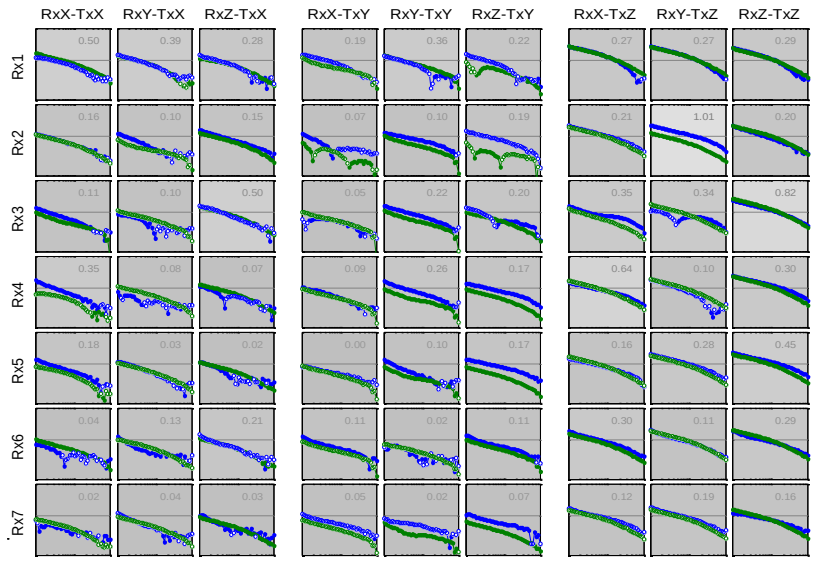


(b) Location of MetalMapper relative to EM61 anomaly

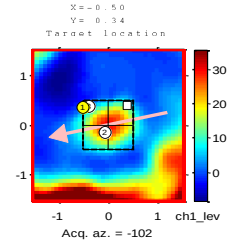


(c) Recovered polarizabilities

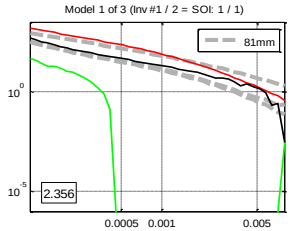
Figure 48: Inversion results for target 2529, a 37mm projectile. An excellent fit with high SNR, recovered positions corresponding to EM61 anomaly location.



(a) Data Fit

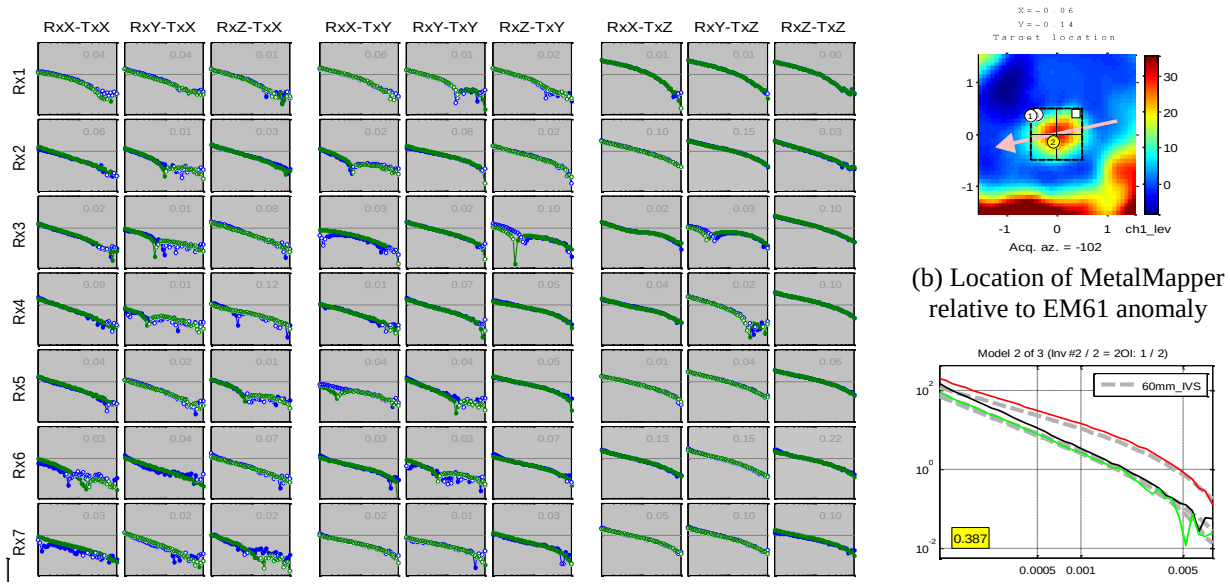


(b) Location of MetalMapper relative to EM61 anomaly



(c) Recovered polarizabilities

Figure 49: Inversion result for target 2136, a 60mm mortar. Results are shown for the single-object inversion. Recovered locations are near the edge of the array at one of the inversion bounds. Polarizabilities are not an excellent match to any of the reference items and one of the secondary polarizabilities is not well constrained. This model was failed as a result (the model from the multi-object inversion corresponding to approximately the same location was also failed for similar reasoning).



(a) Data Fit

(b) Location of MetalMapper relative to EM61 anomaly

(c) Recovered polarizabilities

Figure 50: The QC tool display for target 2136, a 60mm mortar. Results are shown for the multi-object inversion. Recovered locations are in agreement with the EM61 anomaly. Recovered polarizabilities produce an excellent match to the 60mm reference item and this model was passed as a result.

Training Data Selection

Training data was selected using a specialized tool developed for identifying target clusters. A size versus decay feature space plot is generated for all of the passed (SOI and MI) models for each target. Overlain on the same feature space plot are the size and decay features for IVS and test pit items. The user can specify polygons that encapsulate groups of targets expected to be of a specific target type. An example of the Training data tool feature space along with user defined polygons is shown in Figure 51. Once the user has drawn the polygons, the Training data tool can be used to search for models within the defined polygons that have similar polarizabilities. Similar targets will cluster together in feature space. The input parameters (misfit method, time channels used, misfit threshold, cluster size) are all user specified and can be customized depending on the dataset.

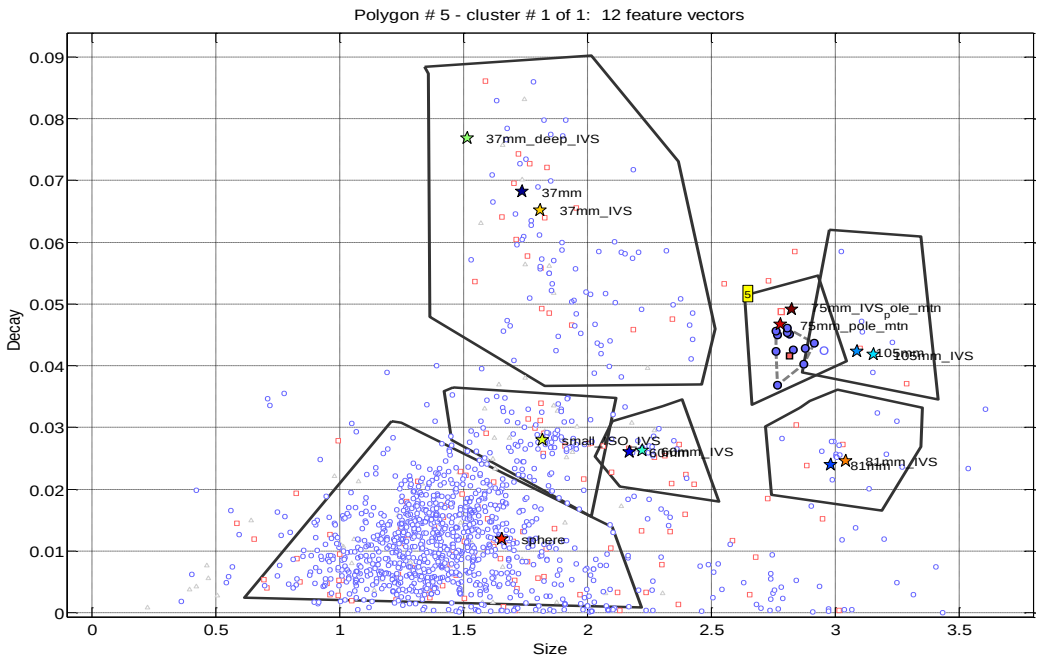


Figure 51: The Training data tool was used to assist the analyst in identifying target clusters in a size/decay feature space.

The first goal in identifying target clusters with the training data tool is confirming that representative items from the IVS and test pit correspond to clusters of similar target types within the feature space of all passed models. Clusters which were found matching the reference polarizabilities of IVS/test pit items were not selected for training data since there was already a high confidence that these items were TOI.

The second important feature of the Training Data tool is the identification of additional target clusters that do not correspond with items in the IVS/test pit data. These additional clusters formed the basis for training data requests. For example, there was a cluster of 24 targets identified which all had three approximately equal polarizabilities (see Figure 52). Two of these items were selected and submitted as part of the initial training data request and both objects were identified as fuzes that should be considered TOI. As a result, a reference polarizability was calculated (the mean of the 24 targets in the cluster) and added to the library of IVS/testpit. The first training data also identified two additional clusters of TOI not present in the IVS/testpit data. A grouping of 60mm mortars clustered differently in feature space than the IVS/testpit data because of the presence of a nose and tail fin which were not present in the IVS/test pit items. There was an additional cluster of 81mm targets identified as well as a non TOI cluster of horseshoes. The first set of training data requested consisted of 8 items (4 TOI, 4 non TOI).

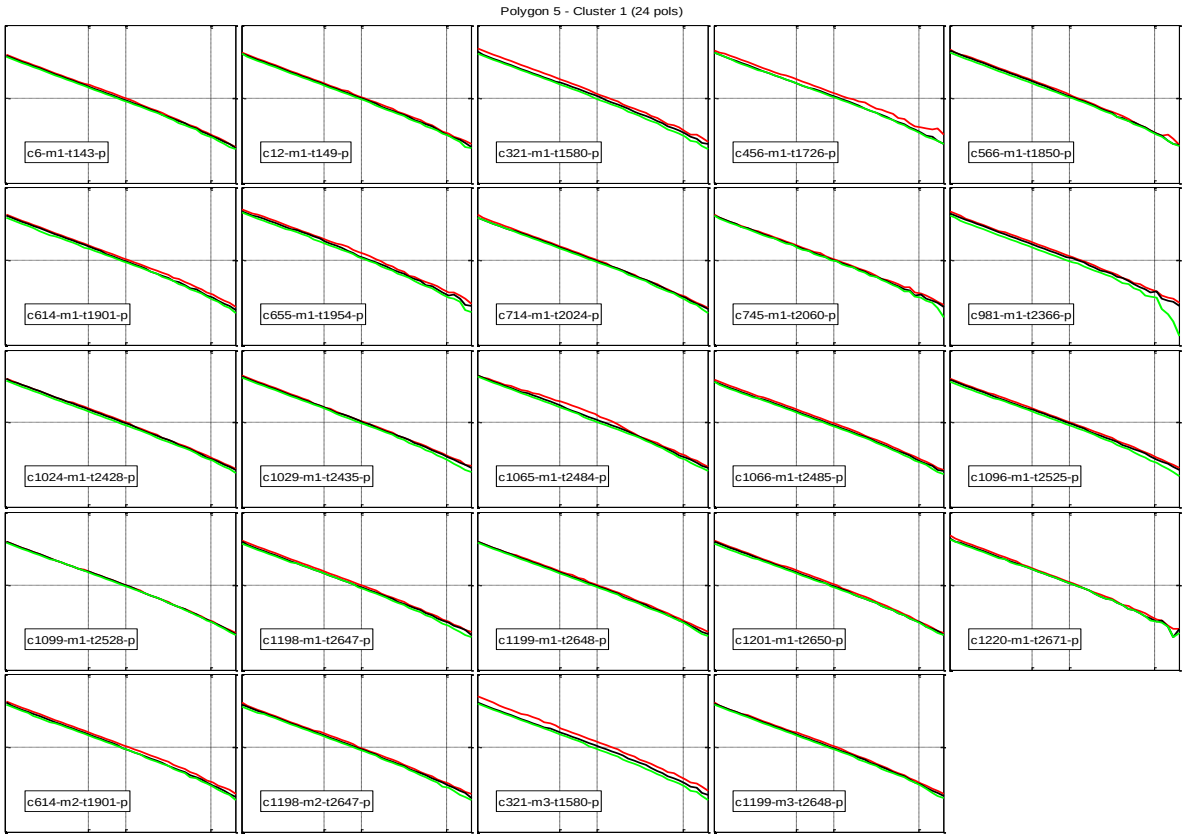


Figure 52: The Training data tool identified this cluster of self similar targets which had 3 approximately equal polarizabilities. Training data requests for these items identified them as fuzes.

A second round of training data was requested. Rather than identifying additional target clusters, the goal of the second round was the investigation of targets that appeared UXO-like but where secondary polarizabilities were not well recovered. A total of 10 additional items were requested of which 6 were TOI and 4 were non TOI. This training data illustrated the need to incorporate a classification strategy that only considered the L1 polarizability after it became evident that some of the TOI training items (see Figure 53) did not have well recovered L2/L3 polarizabilities even though L1 produced an excellent match to the reference item. The challenge of false positives when applying an L1 only classification scheme was also identified (see target 2609 (3 pieces of small frag)) as this item also has an excellent L1 match for a small ISO even though it turned out to be 3 small pieces of frag.

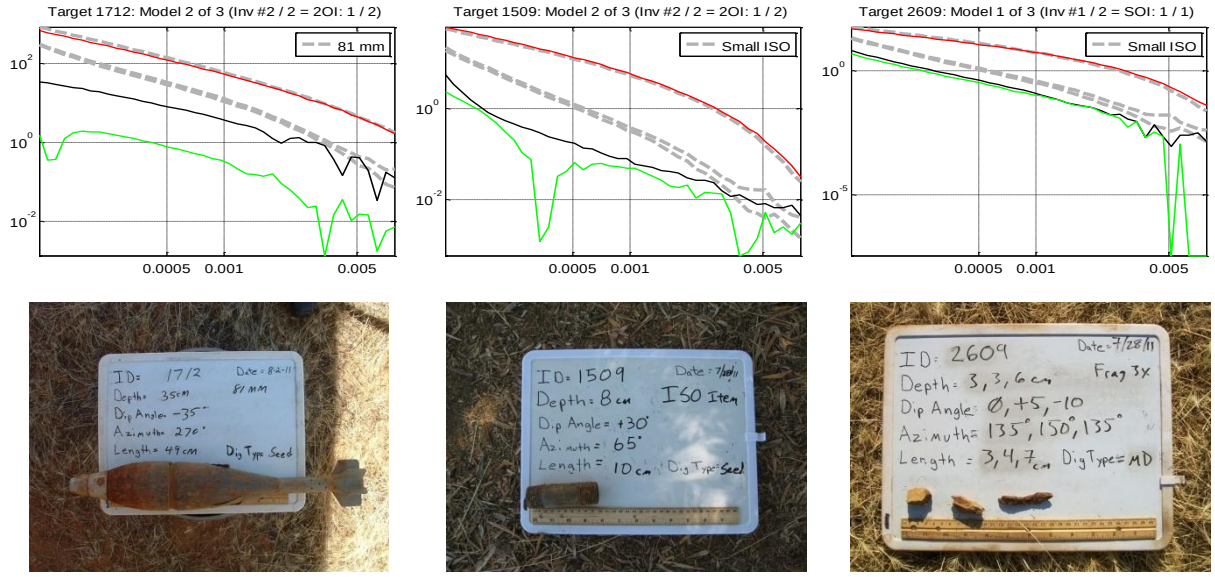


Figure 53: Training data requests for TOI where primary polarizability was an excellent match to reference item however secondary polarizabilities were not well recovered.

Dig List Creation

Another custom tool was designed to aid in the diglist creation process. The tool allows the user to vary a number of parameters that are combined to create a prioritized diglist. The user can specify weightings for polarizability quality, size, decay and misfit method using different combinations of polarizabilities over a range of time channels. Functionality also exists to search the parameter space to automatically select optimal weightings for each parameter. The tool also allows the user to step through the ordered diglist, viewing the polarizabilities along with the best reference item and the location of each item in a size/decay feature space.

A total of four diglists were created for the CH2MHill MetalMapper data. Diglists were prepared using the EM61 data to prescreen out targets that were high confidence not UXO. MetalMapper targets not omitted via the EM61 prescreening, were inverted, analyzed and a prioritized diglist was generated. The EM61 prescreened targets were then appended to the end of MetalMapper diglist as non-TOI items. A full MetalMapper dataset (with no EM61 prescreening) was also analyzed and a diglist generated. For both the prescreened MetalMapper and full MetalMapper datasets two types of classification were applied: one Sky analyst produced diglists based on polarizability matches using the Dig List tool while a second analyst used statistical support vector machine (SVM) classification methods to create two diglists each (for a total of four diglists).

Polarizability Match Diglists

The initial diglist submitted for this method missed three seed items. Target 2277 was a 37mm projectile which was incorrectly included in the prescreened EM61 targets. It was a fast decaying target and its omission was an indication of limitations in classification capabilities when using EM61 data rather than a problem with the process or analyst error (see Figure 27). The two additional missed seeds were targets 1613 (small ISO) and target 1650 (60mm mortar). Both of these were missed due to analyst error. The

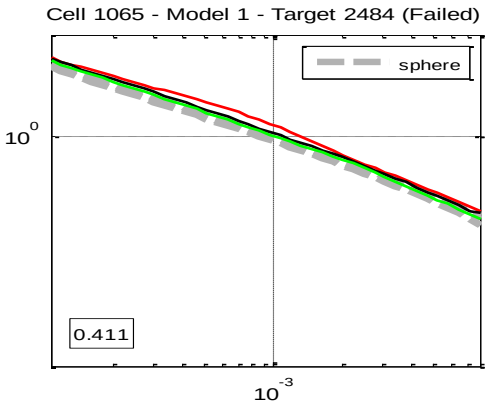
intention was to create a diglist that started sorting by an L1L2L3 polarizability misfit then switching to an L1 only misfit after a number of digs. The analyst using the diglist tool for the first time was unaware that the additional parameters needed to be explicitly set to zero after automatically searching parameter space for optimal values. Both targets would have been identified using the L1 only misfit but failing to exclude the additional parameters (polarizability quality, size, decay) forced these two seeds out beyond the stop dig point because the secondary and tertiary polarizabilities were not well recovered.

Stage 2 Diglist

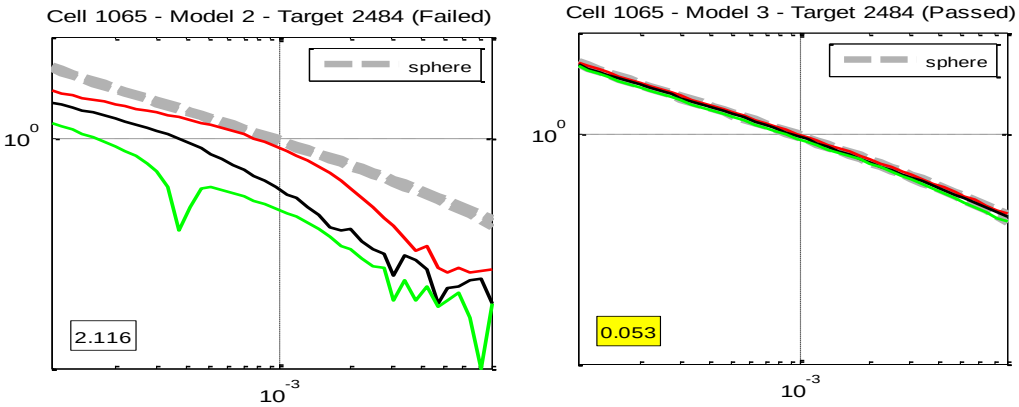
The two missed seeds were classified as training data and the diglist was regenerated as stage 2. At this point the first 200 digs were obtained by ranking targets according to the L1L2L3 misfit. At dig 201, the remaining targets were ranked according to L1 misfit only up to dig 350. The remainder of the digs from 351 onwards were ranked according to a combination of size and decay. Only time channels 1 to 35 (of 42) were used because the late time responses were unreliable and often exhibited large variations from the reference items which tended to push items deeper down the diglist. The stop dig point was set at dig 410 based on a visual inspection of the recovered polarizabilities and the closest reference item. A similar procedure was followed for the all MetalMapper (i.e. no EM-61 prescreening) diglist with the only difference being the stop dig point being set slightly later (dig 426).

Stage 3 Diglist

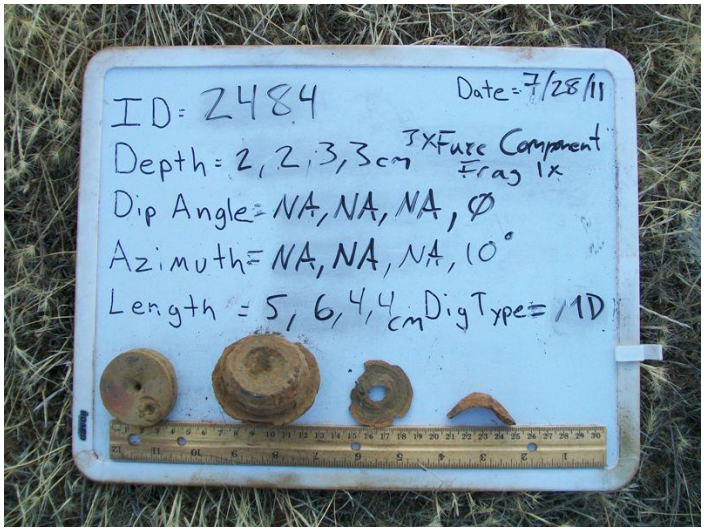
After receiving ground truth information up to the stop digpoint of 410 from Stage 2, two additional targets of interest were noted. A small 5cm fuze not originally identified as a target cluster was revealed in the groundtruth obtained for target 2484 as shown in Figure 54. There was also a larger 6cm fuze (part of the cluster identified in Figure 52) and two pieces of frag recovered at the target location. Initially the second object of the MOI was the only model passed (see Figure 54) as it was a near perfect match for the identified cluster of fuzes. Efforts were made to use the L1 polarization from the first model of the MOI to identify other potential small fuzes since they too were considered TOI. In addition, a number of larger (6cm) fuzes were noted in the groundtruth where not all three recovered polarizabilities matched the reference item. Efforts were also made to identify these large fuzes with only a subset of the 3 recovered polarizabilities matching the reference item. The final diglist submitted was formed by retaining the first 426 digs from stage 2 (up to the stop dig point where groundtruth was received in stage 1). After the stop dig point, potential 5cm fuzes and larger fuzes with only a subset of well recovered polarizabilities were placed then remaining targets were ordered based on decay only to try and identify all slowly decaying targets relatively early in the diglist. Appending these additional suspected fuze items to the diglist moved the stop dig point to dig 488 for the full MetalMapper diglist and dig 446 for the EM61 prescreened diglist.



(a) Single Source inversion result



(b) Two Source Inversion Result



(c) Photo of targets

Figure 54: Groundtruth from the Stage 2 diglist revealed a small 5cm fuze that was not identified earlier through clustering of similar targets in feature space. The L1 polarizability from the multi object inversion was used to search for similar items once the target had been identified.

Results

Although all dig lists were initially scored with fuzes treated as TOI, the program office ultimately decided to treat fuzes as clutter (see Figure 55). There was one TOI item beyond the stop dig point in the submitted dig list. Target 1786 was a small ISO yet there is no indication of that when comparing with polarizabilities from the closest matching reference item (see Figure 57). The polarizabilities for both the SOI and the first model of the MOI resemble a deep 81mm mortar at early times and are not well recovered at later times. Both of these recovered models are near the edge of the sensor and at the X dimension inversion bounds. This item was missed as a result of poor data quality, there is not sufficient information to warrant digging this item earlier in the diglist. The data was not judged to be poor enough quality that the target should be flagged as can't analyzed and dug.. In the future, models that are recovered at any of the inversion bounds should be flagged for closer inspection and potential can't analyze designation.

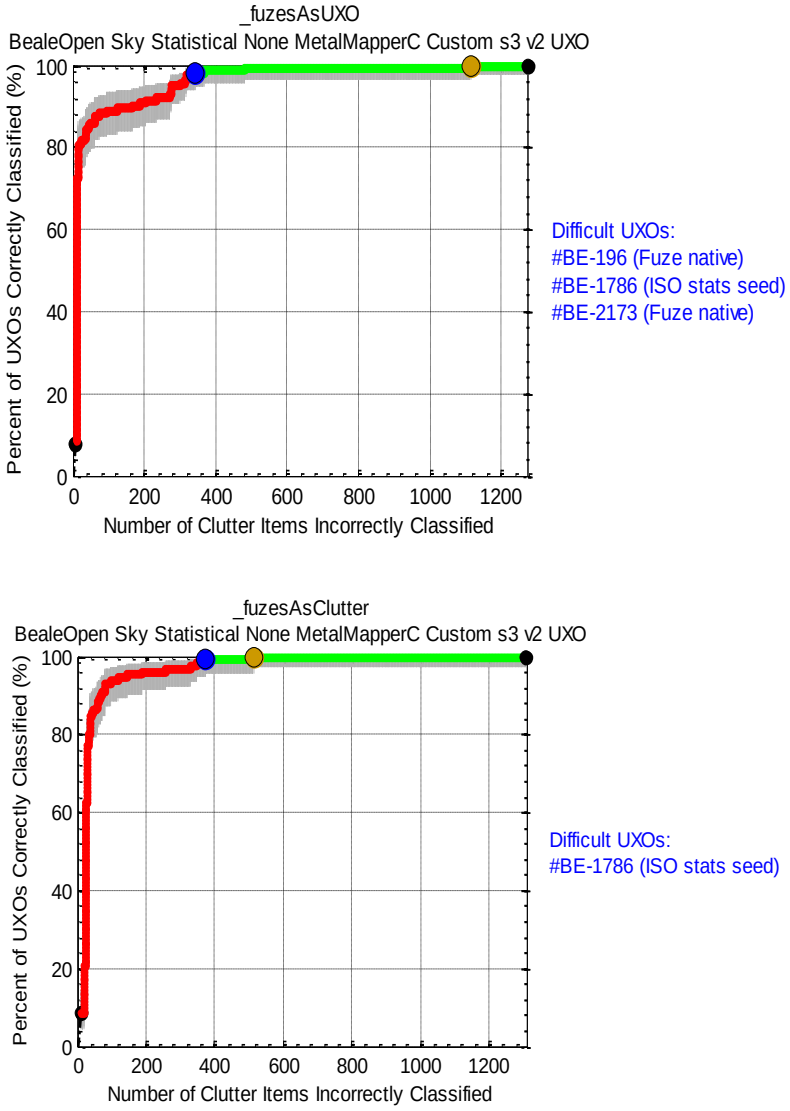


Figure 55: ROC curves for CH2M Hill MetalMapper data for the case where fuzes were considered to be TOI (top) and with fuzes designated clutter (bottom).

If fuzes were considered TOI, two additional items were left beyond the stop dig point (see Figure 56). Target 2173 was 26 digs beyond the stop dig point. Groundtruth indicated that the item was a 6cm fuze yet the recovered polarizabilities seem closer to those observed for the smaller fuzes (although they weren't similar enough to be detected when searching for small fuzes after their discovery in the Stage 2 groundtruth). It also wasn't flagged using the large fuze search that permitted only a subset of the recovered polarizabilities to match the reference item because none of the recovered polarizabilities sufficiently matched the large fuze reference item. Target 196 was a 5cm fuze that occurred at dig 1301 (793 digs beyond the stop dig point). This target produced quickly decaying recovered polarizabilities that were inconsistent with TOI form the site for both single and two object inversions. One of the models produced by the multi object inversion produced a primary polarizability that was a decent match to the fuze library item at early times. Tertiary polarizabilities were not as well recovered. The primary polarizability was a close enough match to the fuze reference values that it would have moved Target 196 significantly higher up the diglist had that model been passed by the analyst. Because the single and multi object inversion both produced models with fast decaying, low amplitude polarizabilities, those were assumed to be the representative polarizabilities and the less well recovered, slower decaying model with a primary polarizability reasonably matching the fuze from the reference library was failed.

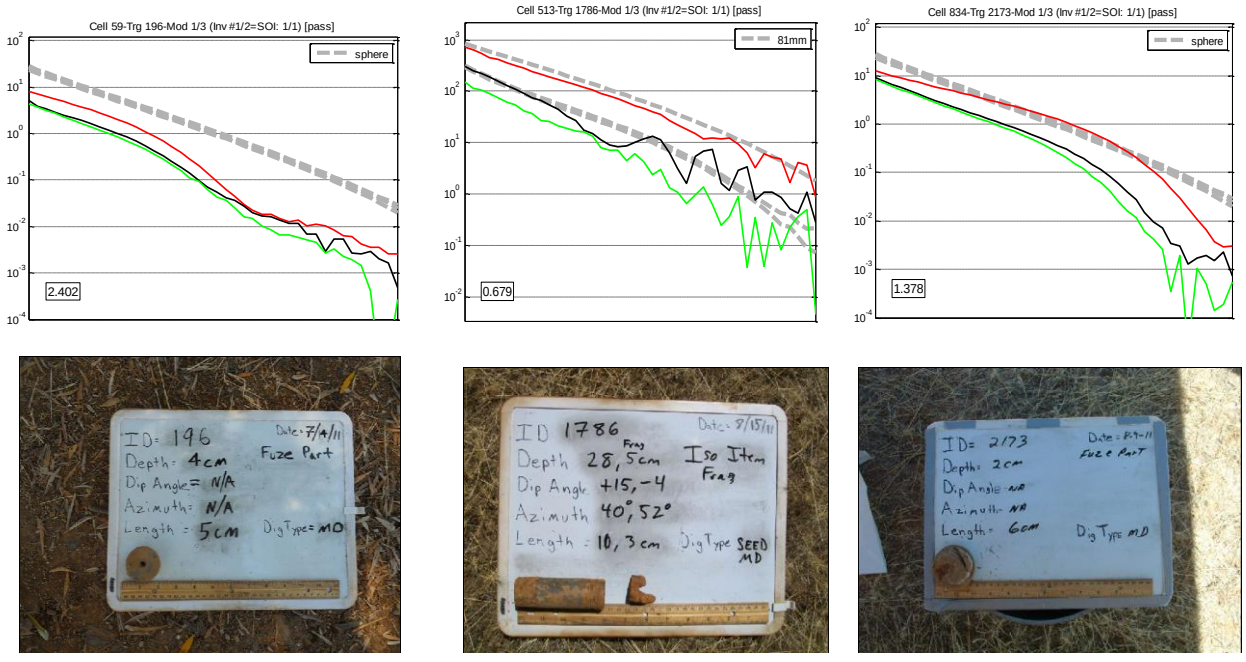
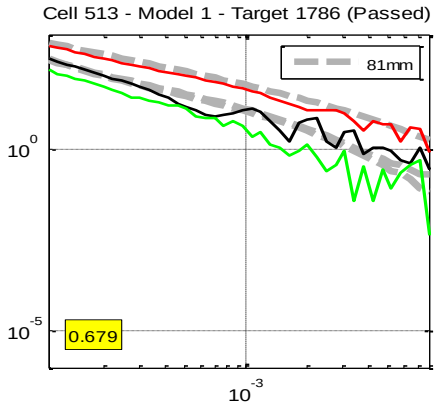
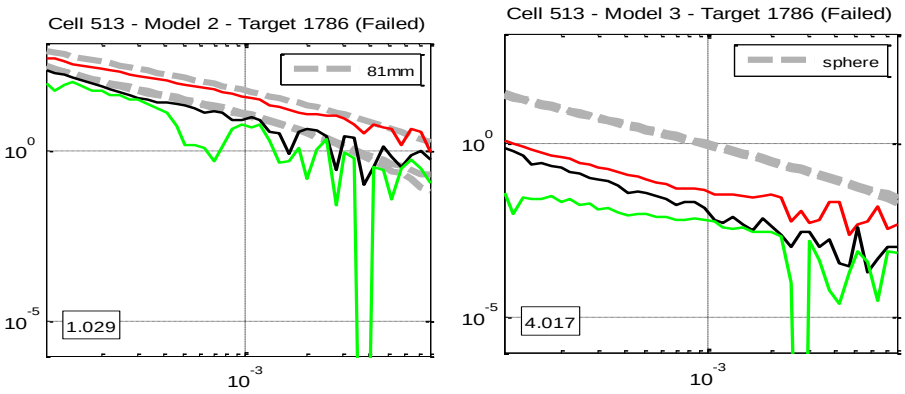


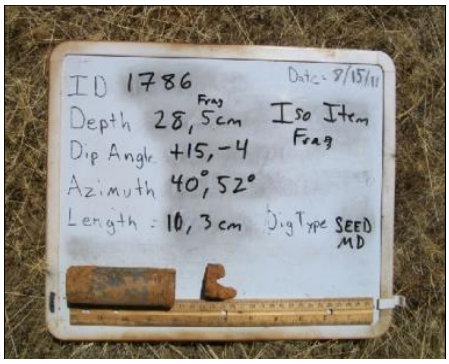
Figure 56: Polarizabilities for the three missed items when fuzes were considered TOI. In addition to the ISO (see Figure 57), 5cm and 6cm fuzes were also not identified as TOI



(a) Single source inversion result



(b) Two source inversion result



(c) Photo of Target 1786

Figure 57: Summary showing key portions of the QC tool for missed ISO (Target 1786). Polarizabilities recovered from both single and multi object inversions do not produce a good match to reference ISO polarizabilities. Target location indicates the recovered location was near the corner of the MetalMapper array at the inversion bounds and should have been flagged as can't analyze.

Appendix C: SUPPORT VECTOR MACHINE DIGLISTS

For comparison with MetalMapper (CH2M and Parsons) diglists generated with analyst intervention, those same data sets were analyzed using support vector machine (SVM) classifiers. For a basic description of the SVM algorithm, see section 2.1.3.

In all cases we relied upon the training requests submitted by the analyst for each individual data set, no additional training requests specific to the SVM classifiers were submitted. The training data for the SVM classifiers therefore comprised TOI features from test pit measurements and training requests. When training statistical classifiers, we do not use non-TOI from training requests as these clutter items are typically queried because they are similar to TOI in size-decay, or polarizability, feature space. Including such clutter items in the training data can lead to poor generalization to the test data as the classifier may overfit training non-TOI that are very close to TOI. We therefore prefer to train SVM classifiers using *assumed* non-TOI identified in the test data as follows:

1. We compute a misfit matrix **M** with elements

$$\boldsymbol{M_{jk}} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} (\boldsymbol{x_i^j} - \boldsymbol{x_i^k})^{\boldsymbol{2}}. \qquad (13)$$

   Here x_i^j denotes the i[th] element of the j[th] training vector, and similarly x_i^k denotes an element of the k[th] test vector. The feature vectors **x** can, in general, be any features derived from the dipole model, here we compute misfits between (log transformed) total polarizabilities

$$\boldsymbol{L_{tot}}(\boldsymbol{t_i}) = \sum_{\boldsymbol{j=1}}^{\boldsymbol{3}} \boldsymbol{L_j}(\boldsymbol{t_i}) \qquad (14)$$

   over the full range of channels for each instrument. Alternatively, this operation can also be carried out using size-decay features, though in this case the features must be normalized so that the size feature does not dominate the misfit.
2. We convert each row of **M** to an integer rank vector, producing the rank matrix **R**. That is, the element M_{jk} with minimal misfit over all elements in the j[th] row is assigned the value 1, while the maximal misfit element, for a total of K test vectors, is assigned the value K.
3. The total rank **T** vector is then the column sum of the rank matrix $T_k = \sum_{j=1}^{J} R_{jk}$
4. Finally, we sort the rank vector in descending order to determine a ranking of test feature vectors that are farthest from training TOI in the sense of the misfit function. The first $N_{clutter}$ items in the sorted list are then used as training non-TOI, with $N_{clutter}$ ranging from approximately 100-500 items, depending on the size of the test data set. Note that in this analysis we do not account for the presence of multiple feature vectors for a given target, so that two passed models from the same item could be used in the set of assumed non-TOI.

Figure 58 shows the training TOI and non-TOI feature vectors identified using this approach for the Camp Beale Parsons data. We emphasize that this analysis is carried out using total polarizabilities, for simplicity we display the selected feature vectors in size-decay space.

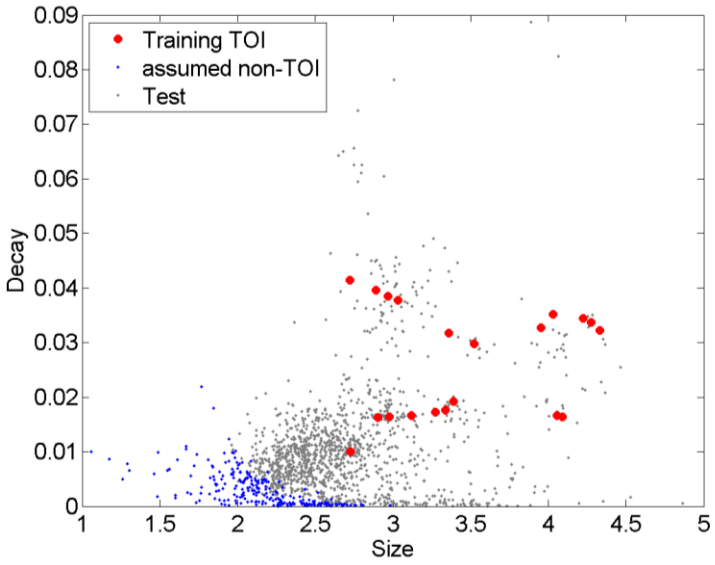


Figure 58. Size-decay space for Beale MetalMapper Parsons data. Assumed non-TOI are test feature vectors with maximum total polarizability misfit with training TOI.

Once the initial training data have been defined, we generate diglists from the outputs of nonlinear SVMs with radial basis functions. We employ a "two-stage" classification approach that combines SVM classifiers trained on two feature sets:

1. All polarizabilities. Secondary and tertiary polarizabilities are more susceptible to noise at late times, and so for these parameters we truncate the time range of channels used for classification. For example, with MetalMapper data sets, we use a longer time range of L_1 (channels 1 (0.11 ms) -35 (3.80 ms)), whereas for L_2 and L_3 we restrict classification to channels 1 to 20 (0.78 ms). Selection of time channels can be automated using feature selection algorithms; however, when dealing with limited TOI training vectors we prefer to use analyst judgment when setting these parameters.
2. Total polarizabilities. The range of channels here is typically the same as that used for L_1, since the total polarizability tends to be dominated by the primary polarizability.

Figure 59 shows a representation of the SVM decision function for a two-stage classifier trained on Parsons MetalMapper data. Two distinct bands of likely TOI regions are apparent in polarizability feature space, corresponding to large (105 mm, 81 mm), and small (37 mm, ISO) ordnance, respectively.

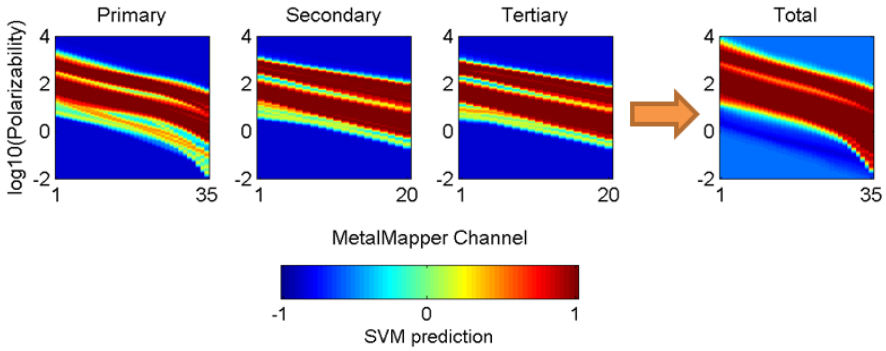


Figure 59. Decision surface for two-stage SVM classifier applied to Parsons MetalMapper data. An SVM prediction of 1 indicates a high likelihood of TOI.

For both feature sets, the polarizabilities are log-transformed prior to training and prediction stages. We use a width of σ=1 for the Gaussian kernels; this choice is based on past applications of this algorithm to polarizability features. Cross validation techniques can also be used to select the kernel width.

Our two-stage approach is motivated by the observation that the majority of targets of interest interrogated with next generation sensors produce well-constrained polarizabilities that are an excellent match to training TOI. In the initial stages of digging we therefore wish to exploit all available information in the recovered model and should use all polarizabilities in classification. However, a small proportion of TOI can still produce poorly constrained transverse polarizabilities, particularly if the sensor is not properly centered relative to the target. In these cases, we find that the total polarizability can still match training vectors. However, in this second stage the false alarm rate will inevitably begin to increase as clutter can often match total polarizabilities for smaller TOI (e.g. 37mm or fuzes). We therefore expect that the ROC for two-stage SVM classification will rise sharply and then "turn over" slightly in the second stage. This strategy was successfully employed for classification of Camp Butner MetalMapper and TEMTADS data. Retrospective analysis on Camp Butner data sets showed that the two-stage approach reduced the significantly final false alarm rate relative to a single SVM that relies on all polarizabilities throughout digging.

When should we make the switch between the first (all polarizabilities) and second (total polarizabilities) stages of digging? A plot of the decision statistic f_{SVM} sorted in descending order has an inflection point where the first stage classifier transitions from clear matches to known TOI to poorer matches (Figure 60). In Figure 60 this corresponds to a value of f_{SVM}≈-0.5, indicating we are digging halfway between the decision boundary (f_{SVM}=0) and the support vectors for the non-TOI class (f_{SVM}=-1). We select this point to transition to the second stage classifier. A similar inflection point is identified in the second stage as our initial stop dig point.

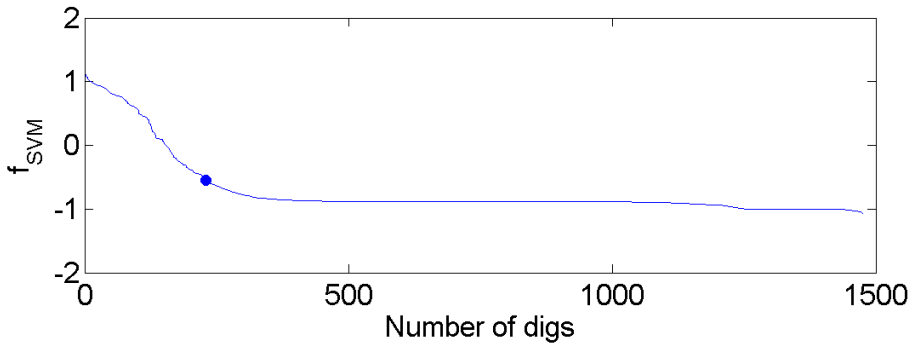


Figure 60. SVM decision statistic f_{SVM} for stage 1 (all polarizabilities) SVM classifier applied to Beale Parsons MetalMapper tests data. Marker indicates point in dig list at which we switch to stage 2 (total polarizability) classifier.

Once groundtruth is received up to our initial stop dig point, we update the SVM diglists by either retraining with newly encountered TOI, or, if necessary, extending the stop dig point to achieve a specified confidence that no more TOI remain in the ground. We describe these procedures in the following sections.

Selecting a stop dig point

If no novel TOI are encountered up to the current stop dig point, then we have some confidence that the training data have adequately characterized the test data and no TOI will be left in the ground with the current classification strategy. To formalize this determination, we use a binormal model of the ROC to arrive at a final stop dig point. This model assumes that the observed ROC can be represented as a sample from two normally-distributed score distributions. The resulting ROC curve is a function of two parameters (Metz et al, 1998)

$$\boldsymbol{a = \frac{|\mu_1 - \mu_2|}{\sigma_1}} \qquad b = \frac{\sigma_1}{\sigma_2}. \tag{15}$$

Maximum likelihood estimation of these parameters yields the predicted ROC and the model covariance, which can then be used to determine confidence intervals on the ROC. In practice, the generating score distributions of TOI and non-TOI are rarely normally distributed, and so estimating means and variances directly from the empirical score distributions will yield a poor fit to the observed ROC. A better strategy is to express the predicted ROC as a function of the parameters *a* and *b* and to fit the observed ROC directly. This approach can often yield an excellent fit, even if the underlying score distributions are not normally distributed (Hanley, 1988). This is because the observed ROC is invariant under arbitrary monotonic transformations of the decision statistic (i.e. transformations that preserve the ordering of the diglist).

We use this model to test whether all ordnance have been found at a selected dig point as follows:

1. Fit a binormal model to known ground truth (i.e. up to selected stop dig point), assuming as our null hypothesis that the number of detected targets of interest N_{TOI} at the site is the number of TOI found thus far.
2. Determine a confidence interval on the expected false alarm rate by finding the point at which the estimated binormal confidence interval has the value

$$\boldsymbol{P = 1 - 1/N_{TOI}}. \tag{16}$$

For the data in Figure 61 this interval is shown as a solid horizontal line.

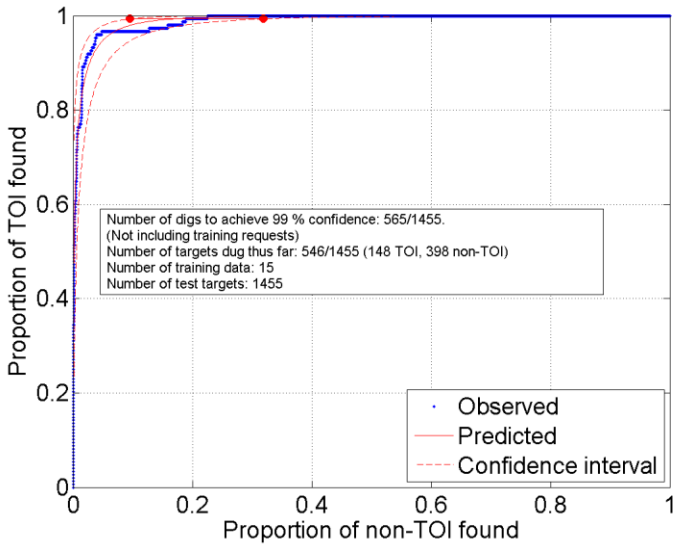


Figure 61. Predicted binormal ROC and 99% confidence interval for first stage groundtruth from Beale Parsons MetalMapper data. An additional 19 digs (from 546 to 565) are required to test the null hypothesis.

We then test our null hypothesis at the specified confidence by digging out to the maximum extent of the confidence interval on the FAR. If no further TOI are encountered then we retain the null hypothesis and finish digging. If TOI *are* encountered then we retrain our classifier, as described below.

Retraining classifiers

When new TOI are found during digging we add them to our training data set and retrain our two stage SVM classifier. Newly labeled non-TOI are not included in the training data, throughout the classification process we rely on assumed, fast-decaying test vectors. The reordered diglist will ideally move any test items that are similar to new training vectors earlier in the dig order. For Parsons MetalMapper data, one new TOI was encountered after digging out to initial 99% confidence interval, as described in the previous section. Figure 62(a) compares the target ordering before and after retraining with this new TOI added to the training data. While some items have been bumped up the dig order, as desired, retraining has an unintended effect of significantly moving other targets later in the dig order. Any reordering will of course affect the ranking of all remaining test targets, but we do not want retraining to move any targets significantly later than in our initial rankings.

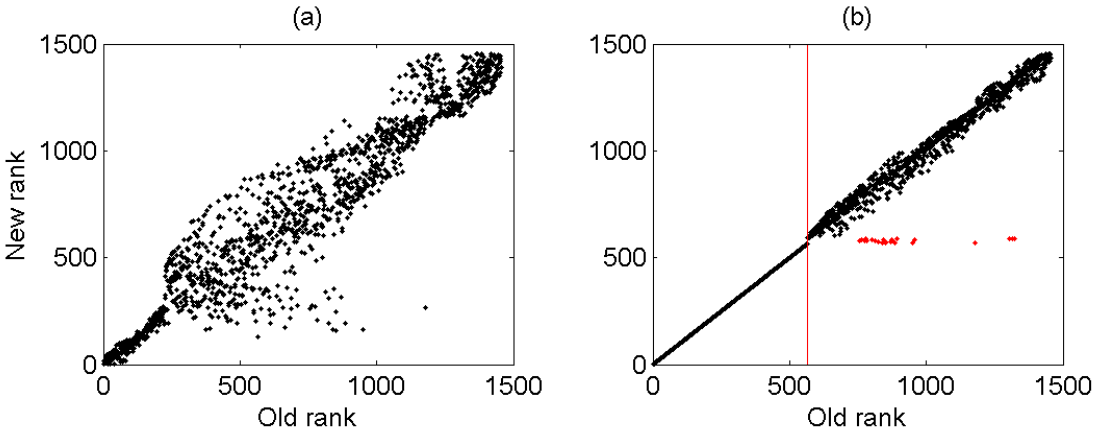


Figure 62 (a) Effect of retraining on dig order for full test data set (b) Final dig order after re-ranking to minimize undesired increases in rank. No re-ranking is applied for the first 565 digs (indicated by vertical red line) because these items have already been labeled in the first stage of digging. Unlabeled targets bumped to the start of the dig order are shown in red.

To minimize movement of targets down the diglist during retraining, we have developed a re-ranking algorithm that compares the new and old (prior to retraining) ranks of a target and uses the minimum of the two in the final list (here a small rank denotes an earlier dig). In addition, we identify a number N_{bump} of unlabeled targets with the largest change in rank following retraining (i.e. the targets that move most up the diglist). We then bump these items up to the start of test items in the final, revised diglist. Figure 62(b) shows the final diglist and the "bumped" targets after this re-ranking has been applied. The first part of the diglist is unchanged because these items have already been labeled and so no reordering can be applied. For the remaining items there is a much stronger correlation between the final ranking and the old ranking, with minimal movement of targets down the diglist (i.e. above the 1-1 line).

Retrospective analysis of SVM classification at Camp Beale

Figure 63 compares classification results for support vector machines applied to the Parsons MetalMapper data. In this example only training TOI from initial requests are used to train the classifiers, we do not include information about native fuzes that was obtained from partial ROC curves. We consider SVM classifiers trained on

1. All polarizabilities, with secondary and tertiary polarizabilities limited to the first 20 time channels,
2. Total polarizability, using all 42 channels,
3. A two-stage SVM classifier, as described on page C-2. This was our first submitted diglist.

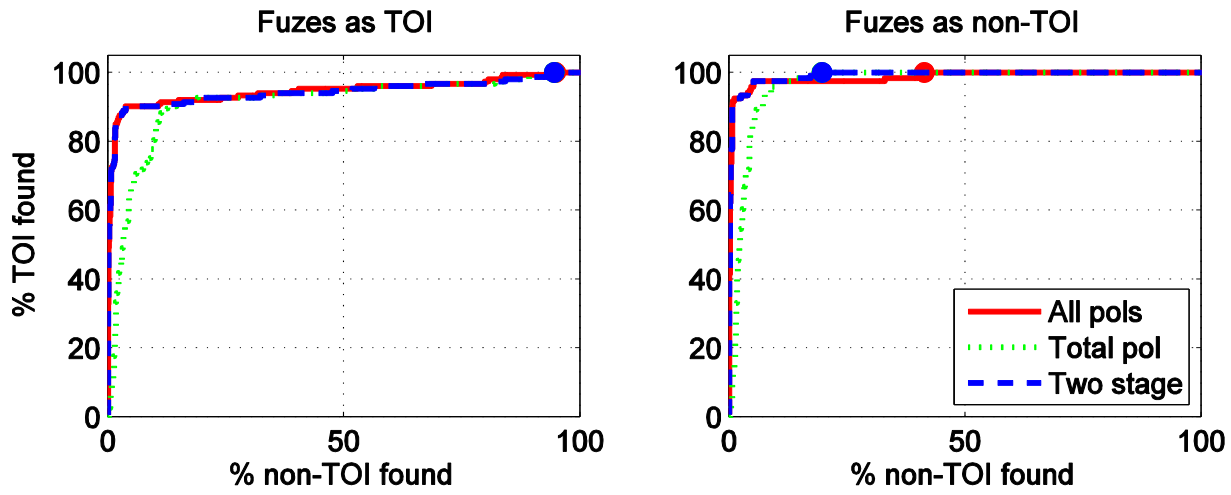


Figure 63. Retrospective comparison of SVM classification strategies applied to Camp Beale Parsons MetalMapper data.

For classification of large TOI (i.e. treating fuzes as clutter), the two-stage algorithm has the advantages of both the "all polarizability" classifier (good initial performance on high SNR TOI), and the "total polarizability" classifier (good classification of low SNR TOI with poorly constrained secondary polarizabilities). This is consistent with results obtained at the Camp Butner demonstration. The final false alarm rate of the two-stage approach is identical to that obtained if we just use total polarizabilities throughout classification. Our method for selecting the switchover between the two stages seems to be effective: the two-stage ROC departs from the "all polarizability" ROC just as the latter starts digging primarily clutter.

None of the methods considered here has any advantage when finding small fuzes. We maintain that any automated classification method will fail to find the unique, small fuzes encountered at this site. Expert analyst input must be incorporated into classification in order to identify these items.

Appendix D: PROCESSING AND CLASSIFICATION ANALYSIS OF BUD HAND-HELD DATA AT CAMP BEALE

This report summarizes the processing of BUD (Berkeley UXO Discriminator) Hand-Held Data at Camp Beale. Multiple BUD soundings were collected at each flagged anomaly. We chose to invert two single sounding data for each anomaly: one data set for the flag location and one selected with maximum ZZ response. The inversions were run with single- and two-object models. Visual QC was performed in conjunction with several metrics of misfit, data SNR, and recovered polarizabilities. For this data set, we mainly used a library matching method to perform training data selection and the ranking of anomalies. Based upon library matching, we partitioned test data into 6 sets. In retrospective analysis, we found these heuristic partitions were reasonable and useful to catch high confidence TOIs, low confidence TOIs, and clutter sets. The main challenge remains with the low confidence TOI sets from which the majority of false positive anomalies occur. Given the high 98.93% PPV (positive predictive value) of the high confidence TOI set, we showed that the diglist ordering benefits from incorporating information (such as distribution of polarizability strength, target depth and range) contained in the high confidence TOI set into the low confidence sets. The retrospective experiment did avoid a number of unnecessary digs, improving the classification results.

BUD Hand-Held Instrument and Test-Pit Data

The BUD (Berkeley UXO discriminator) hand-held TEM system consists of three orthogonal transmitters and ten pairs of dB/dt induction receivers that measure gradient fields. The use of differential modes is expected to reduce the ambient and motion noise, and produce enhanced sensitivity to the gradients of the target. The footprint of the instrument is 35 cm. The system has 46 logarithmically spaced time channels ranging from 0.08 ms to 1.46 ms. A single sounding location is comprised of 30 x 46 spatial-temporal measurements. Figure 64shows the sensor.

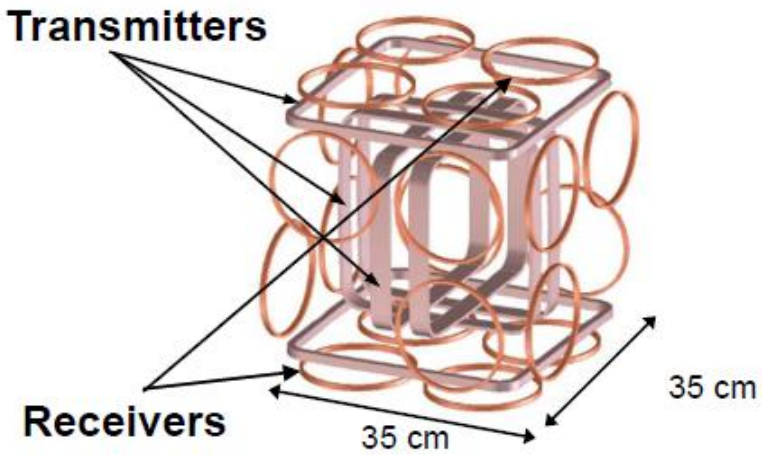


Figure 64: Schematic of the HandHeld BUD

To evaluate the instrument, test-pit data were collected using a 37 mm projectile, 60mm mortar, 81mm mortar, and 105 mm projectile/HEAT shown in Figure 65. These items represented the suspected range of munitions in the demonstration area at Camp Beale. The recovered polarizabilities for each item are presented in Figure 66. for various depths and orientations (horizontal, vertical (nose up and down), tilted at 45^{o}). For small objects, recovered polarizabilities generally exhibit small variations (Figure 66a-b). For larger UXOs like 81mm mortars and 105mm projectiles, the recovered polarizabilities displayed significant variations, particularly when the item was oriented nose down and the tail portion of the object is the dominant response measured by the sensor. Impacts of higher order (non-dipole) terms may become significant in those large object cases.

Field Data at Camp Beale, Inversion, and QC

The system was operated in a cued mode. For each anomaly, there are 5 or 3 associated data sets that were collected at the marked locations (flags), and locations at ±0.15 m away from flags (left and right and/or up and down, see Figure 67). This process of collecting multiple locations per flag was beneficial for cases where the target was not directly under the flagged location. By sampling surrounding locations, the instrument was brought closer to the off-flag anomaly and therefore higher SNR TEM responses were achieved.

We inverted single sounding data as follows:

1. Automatically select a sounding that is closest to the target anomaly. This selection is determined by the maximum ZZ response. ZZ responses are referred to the z-receiver responses recorded after firing the z-transmitter.

2. Flag sounding location

3. 2-object inversion (2OI) on the automatically selected sounding data.

4. Invert anomalies at other sounding locations in the QC process if necessary.

These three kinds of inversions were run through UXOLab. Several criteria were used in the QC process: data SNR, data fit error, and the visual appearance of the polarizability curves. A total of 21 bad fits or noisy data (low SNR) were identified marked as "fail". These anomalies were put into the category of "cannot be analyzed" in the subsequent analysis. Groundtruth revealed that these failed results correspond to 14 cases with no contacts, 5 cases munitions debris , 1 case 30 cal, and 1 case 81 mm mortar (buried depth of 47cm). In addition, there are 4 data sets (BE-609, BE-823, BE-879, BE-887) unavailable in the BUD data collection.

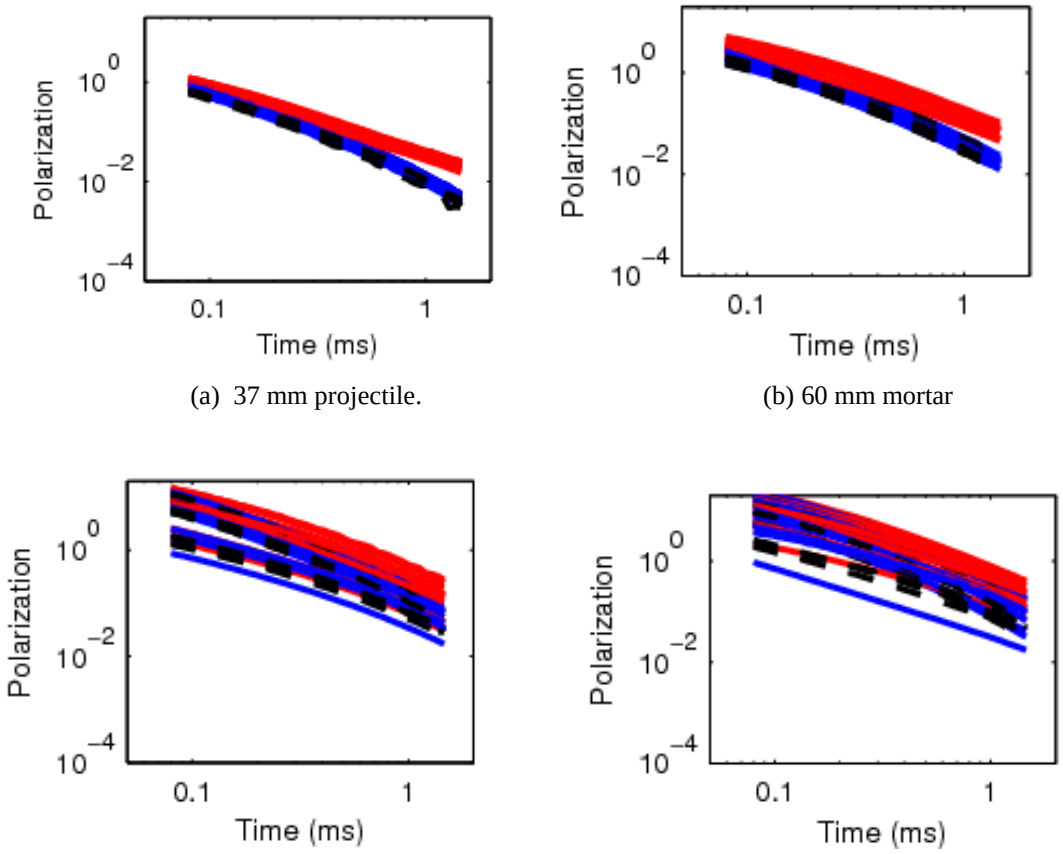
(a) 37 mm projectile

(b) 60 mm mortar

(c) 81 mm mortar

(d) 105 mm HEAT

Figure 65: Ordnance of interest: (a) 37 mm projectile. (b) 60 mm mortar. (c) 81 mm mortar. (d) 105 mm HEAT.



(a) 37 mm projectile. (b) 60 mm mortar

(c) 81 mm projectile. (d) 105 mm mortar

Figure 66: Recovered polarizations from the test-pit data. (a) 37 mm projectile. (b) 60 mm mortar. (c) 81 mm mortar. (d) 105mm projectile. In each plot, the same item was positioned at various depths and orientations (horizontal, vertical (nose up and down), tilted at 45^{0}).

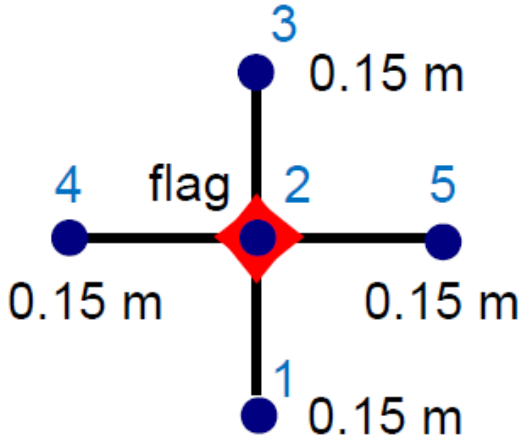


Figure 67: Measurement template in Camp Beale.

Training data selection

To request training data from the test site, we first examined the results by plotting features of polarizability magnitude (size) and decay defined as:

$$\begin{aligned} size\ (t_j) &= \sqrt{\sum_{i=1}^{3} L_i^2\left(t_j\right)} \\ decay\ &= \frac{size\ (t_k)}{size\ (t_j)} \end{aligned} \quad , \tag{6}$$

where L_i (t_j) is the ith principal polarizability component at time channel t_j and $t_j < t_k$. Figure 68 shows the 2D feature distribution. A visual inspection indicated two 'suspicious' clusters (dashed ovals) that are different from the ordnance features of 37mm projectiles, 60mm mortars, 81 mm projectiles, and 105 mm projectiles. Further inspection of the full polarizabilities within the two clusters revealed that the two groups are likely associated with a rod-like object (the top oval cluster) and a sphere-like object (the bottom oval cluster). Therefore, the representative data in these two clusters were selected as part of training sets. Training requests subsequently revealed these clusters as ISO (industry standard object) and fuzes. In Figure 68, we attempted to map the minimum (mfp) misfit of polarizabilities recovered from the test data to the references (including ISO and fuze) into the size-decay feature space and have three rough partitions of the 2D feature space: mfp ≤ 0.5, orange 0.5 < mfp ≤ 1 and green mfp > 1.0, for an initial analysis

The above rule was used to select training data from visually anomalous clusters. We also selected training data according to misfits between test and reference polarizabilities as follows:

1) grouping a set of anomalies whose primary and secondary polarizabilities are close to reference polarizabilities;
2) grouping a set of anomalies whose primary polarizability is close to a reference primary polarizability;
3) grouping a set of anomalies whose secondary and tertiary polarizabilities are close to reference polarizabilities.

We thought that these anomalies might pose difficulties in the classifications stage and thus require attention during the training. As a result, we requested a small number of requested training data: 36. Among these requested data, there were 8 TOIs including 4 fuzes, one ISO, one 37 mm projectile, two 81 mm mortars, 28 munitions debris and horseshoe. Figure 69 presents the recovered polarizabilities for several training data. BE-680 in the top row and BE-905, BE-798, BE-143 in the middle row were selected from the two suspicious clusters identified in Figure 68 and are actually ISO and fuze items. Without ground truth of the training data, they had appeared close to the polarizabilities of 37 mm projectiles. The anomaly BE-411 was requested as it was close to 60 mm mortar but ground truth revealed it was a 37 mm projectile. BE-355 and BE-903 in the middle row were selected because their polarizabilities partly overlapped with those of the105 mm projectile. Ground truth indicated both anomalies were 81mm mortars.

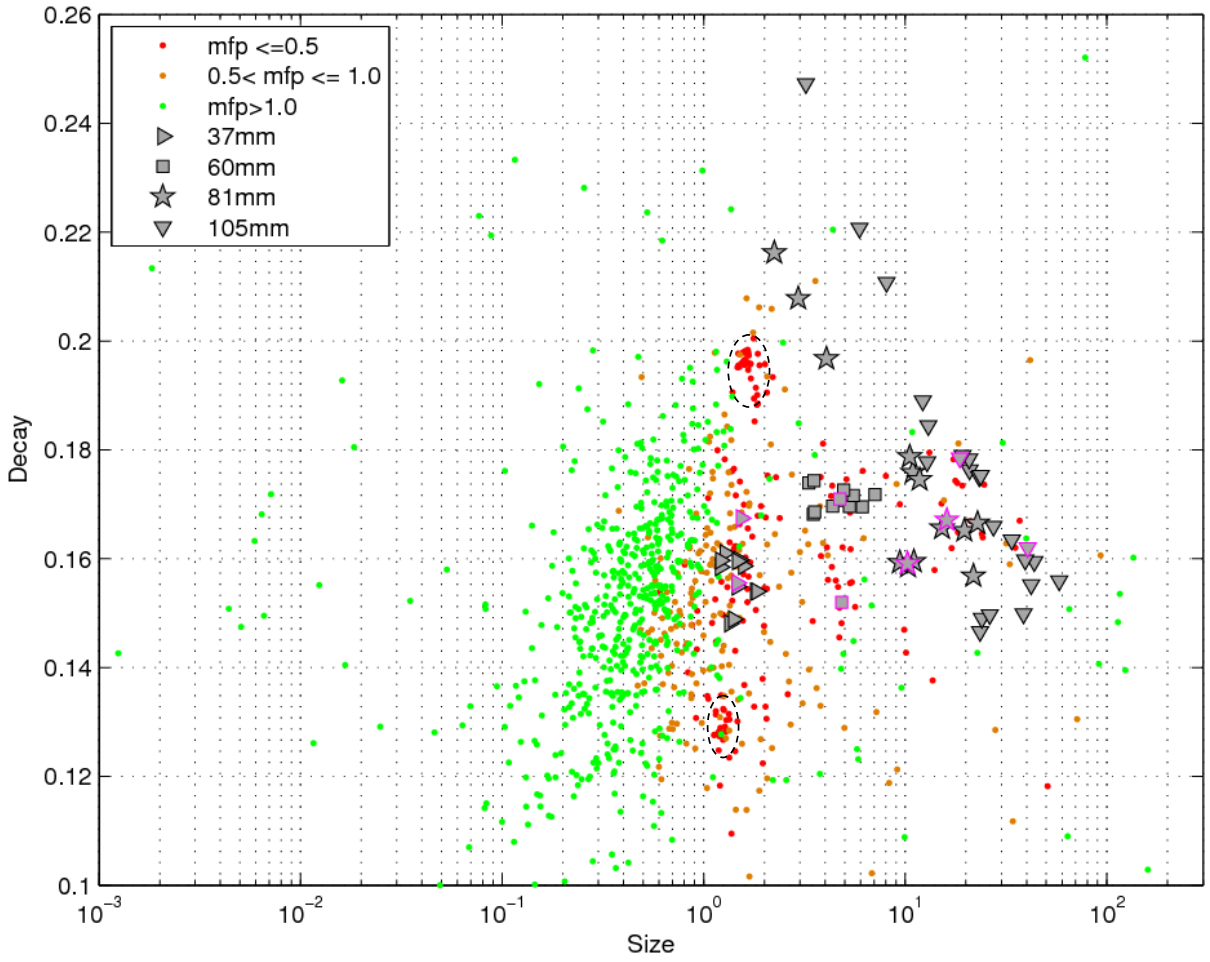


Figure 68: Scatter plot of size and decay for BUD data sets at Camp Beale. Based on the minimum (mfp) misfit of polarizabilities recovered from the test data to the references., we have the 3 rough partitions of the data points for an initial analysis. There are the two potential clusters (dashed ovals) visually observed and are isolated from the ordnance features of 37mm projectiles, 60mm mortars, 81 mm projectiles, and 105 mm projectiles.

BE-669 in the middle row was requested since it resembled an 81mm mortar however ground truth results showed BE-669 was a fuze. The fuze polarizabilities of this anomaly are recovered in the two-object inversion shown in Figure 70. The bottom row in Figure 69 shows four anomalies including: BE-301 (munitions debris), BE-458 (horseshoe), BE-378 (frag), BE-367 (frag). In spite of the fact that none of these items are TOI, they exhibited a high degree of similarity to 37mm projectiles and 81mm mortars. Training data selection serves to expand knowledge of the dataset under consideration. ISO and fuzes discovered via training data were added to our polarizability library and some challenging TOI anomalies were singled out. Knowledge obtained via training data requests is then incorporated into the classification approach.

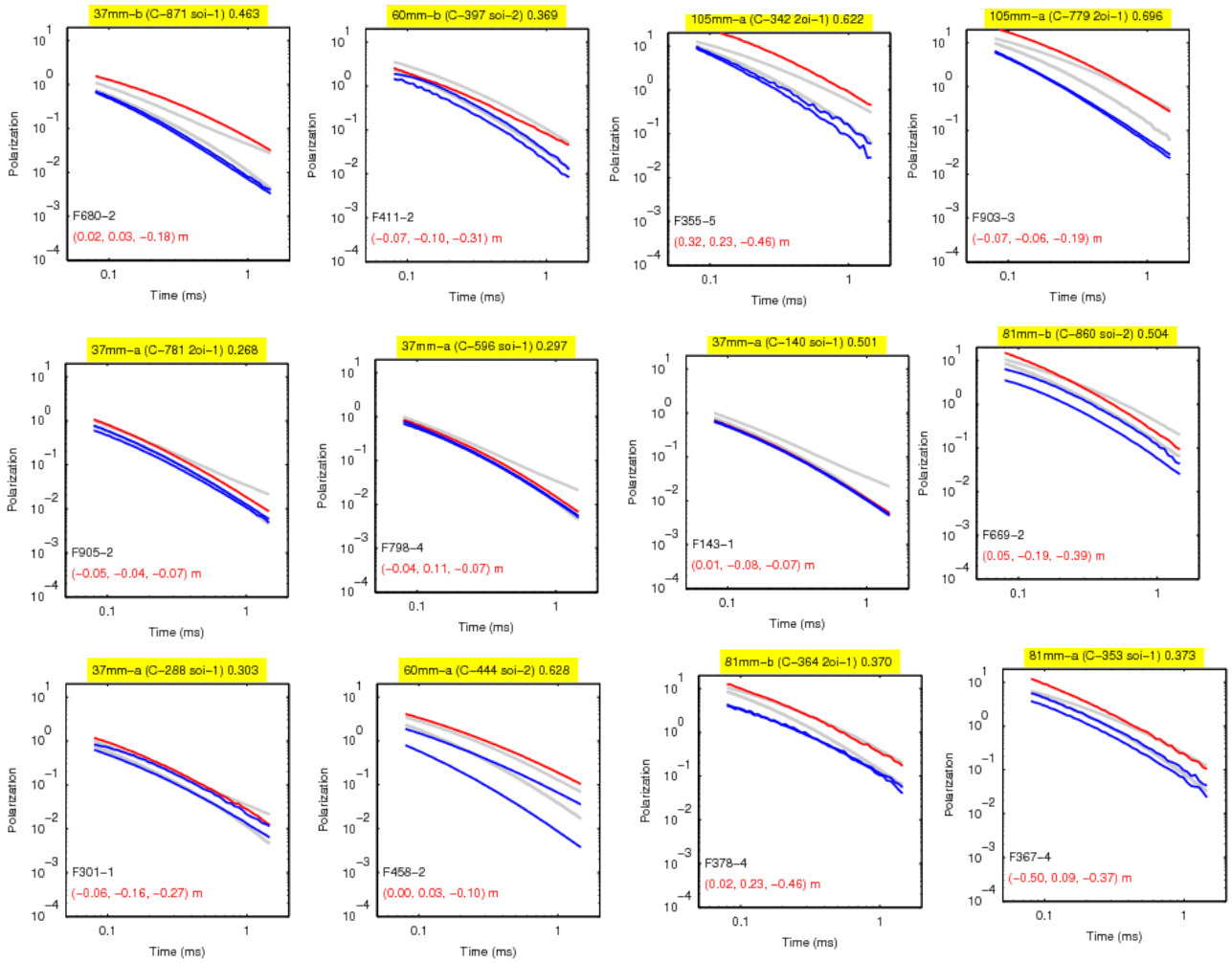


Figure 69: Recovered polarizations for the training data. Top row: BE-680 (ISO), BE-411 (37mm projectile), BE-355 (81mm mortar), BE-903 (81mm mortar). Middle row: BE-905 (fuze), BE-798 (fuze), BE-143(fuze), BE-669 (fuze). Bottom row: BE-301 (munitions debris), BE-458 (horseshoe), BE-378 (frag), BE-367 (frag).

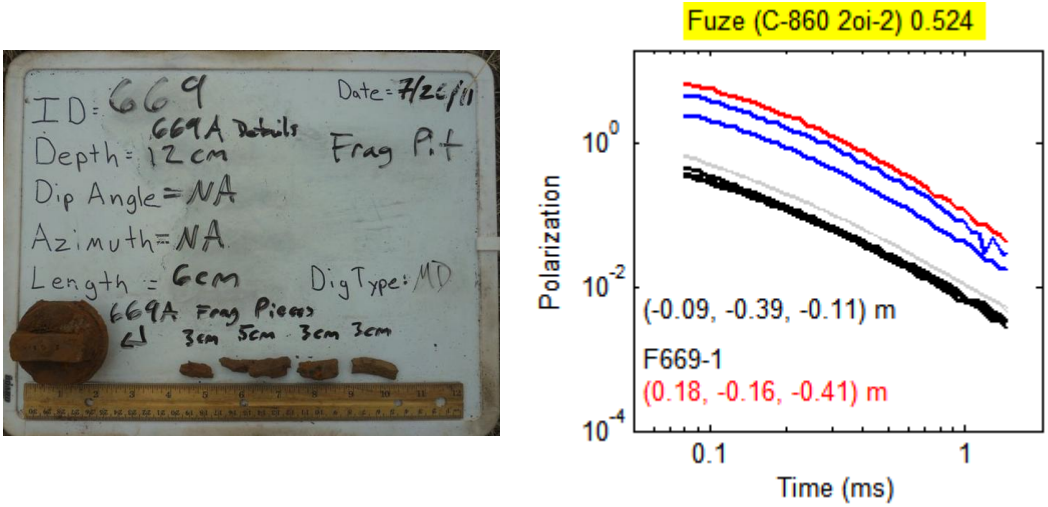


Figure 70: BE-669. The Ground truth and the recovered polarizabilities from a two-object inversion (2OI).

Manual training data selection was tedious and can be inconsistent even with experienced analysts. Better tools are required to guide (semi)-automatic training data selection in the future.
After the first training request revealed two additional new TOIs: ISO (rod-like) and fuze (sphere-like), it was deemed necessary to consider another round of training data requests . In particular, the small nature of these targets and the potential overlap with clutter items suggested additional attention was required.

Ranking anomalies and classification

With information gleaned from training data, we had an updated polarizability library that included both ISOs and fuzes. The updated library, which had 6 targets of interest: fuze, ISO, 37mm projectile, 60mm mortar, 81 mm mortar, and 105mm projectile, was used to rank and classify the test data. To rank the anomalies, a simple library matching method was used. The method computes the fit of a set of test polarizabilities to all 6 sets of reference polarizabilities and selects the minimum fit value (mfp) as a rank value for that test anomaly. Using visual inspection of the correlation between polarizability curves and the mfp values, we partitioned the test anomalies into 6 groups:

Stage 1:

1) *H-TOI set*

This set of anomalies had a high confidence to be ranked as TOIs. Primary, secondary and tertiary polarizabilities fit very well with the references and with mfp ≤ 0.255. It resulted in the total of 94 items identified to be dug. Of those 94, It turned out 93 were TOIs including: 18 fuzes, 22 ISOs, 24 37mm projectiles, 14 60mm mortars, 7 81mm mortars, 8 105mm projectiles. That false alarm anomaly was ranked as a 60 mm item, BE-538. Figure 71 shows that the recovered polarizabilities agree well with the reference 60 mm mortar in addition to a good data fit. However, the ground truth is a piece of frag buried only 4 cm. This is a typical example where a small and shallow object can be modeled as a large object as if it were buried deeply and/or around the sensor edge.

2) *L-TOI set*

In this set, we ranked these anomalies as likely TOIs with $0.258 \leq$ mfp ≤ 0.483. The set has 54 items. Among them, we selected 44 additional items to be dug by referring to training data and considering polarizability characteristics (two equal primary polarizabilities, fast decay, large depth). Groundtruth results identified 13 TOIs and 31 false positive items.

3) *LL-TOI:*

This is a set of low likelihood (LL) TOI with $0.501 \leq$ mfp ≤ 0.657, comprised of 57 items. Using the same selection process as in the L-TOI set, we selected 29 more items to dig and found only one TOI (BE-4 (81 mm mortar) caught as a 60mm mortar)

4) *LVL-TOI:*

This set refers anomalies categorized as likelihood very low (LVL) TOI with $0.659 \leq$ mfp ≤ 0.833, having 50 items. Among them, 24 were ranked as items to dig and all turned out to be false positives.

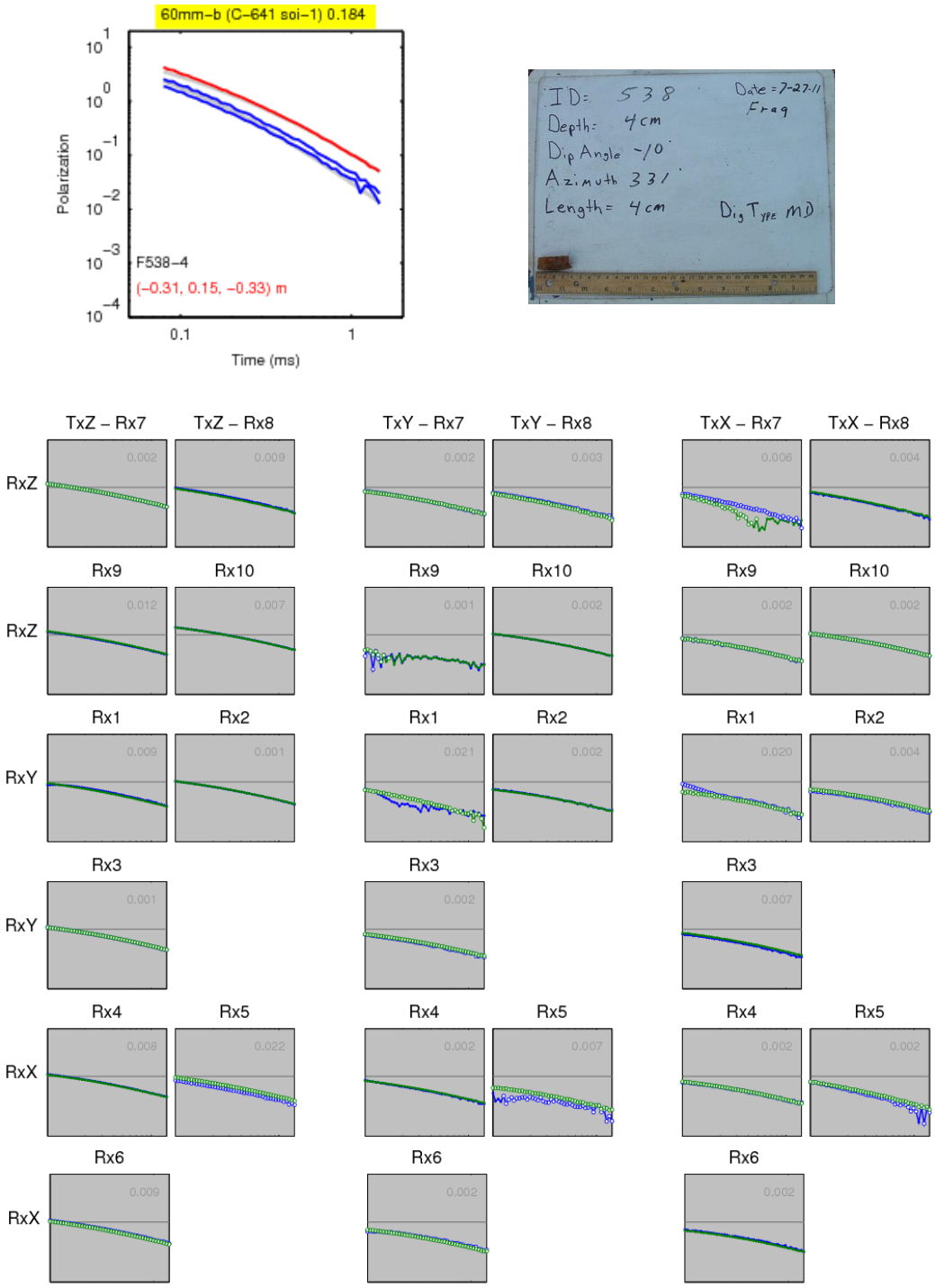


Figure 71: BE-538. The recovered polarizabilities, the ground truth picture, and the fits between the observed (green) and predicted (blue) data.

5) *L-Clutter*

This is the set of anomalies that were most likely clutter and divided into 4 subsets

5-1) subset with 0.835 ≤ mfp ≤ 0.953. 48 items. 28 items were marked to be dug including one TOI item (ISO).
5-2) subset with 0.962 ≤ mfp ≤ 1.101
5-3) subset with 1.104 ≤ mfp ≤ 1.266
5-4) subset with1.272 ≤ mfp ≤ 1.405

6) *H-Clutter*

The remaining anomalies were considered to be "highly possible clutter" and were not included in the digging analysis.

At stage 1, we dug 219 anomalies and found 108 TOIs. Results of the stage-1 dig list are summarized in Table 9.

Table 9: Results of stage-1 dig list. PPV: Positive predictive value

Set	Dug	TOI	Frag	PPV
H-TOI	94	93	1	98.93%
L-TOI	44	13	31	29.54%
LL-TOI	29	1	28	3.44%
LVL-TOI	24	0	24	0.00%
L-Clutter	28	1	27	3.57%
Total	219	108	111	49.31%

At stage 2 and 3, we reexamined the first 5 sets of test anomalies and selected an additional 12 anomalies to be dug, finding one TOI item in the set of LL-TOI. That is BE-854 caught as likely an ISO but ground truth revealed to actually be frag + fuze (Figure 72).

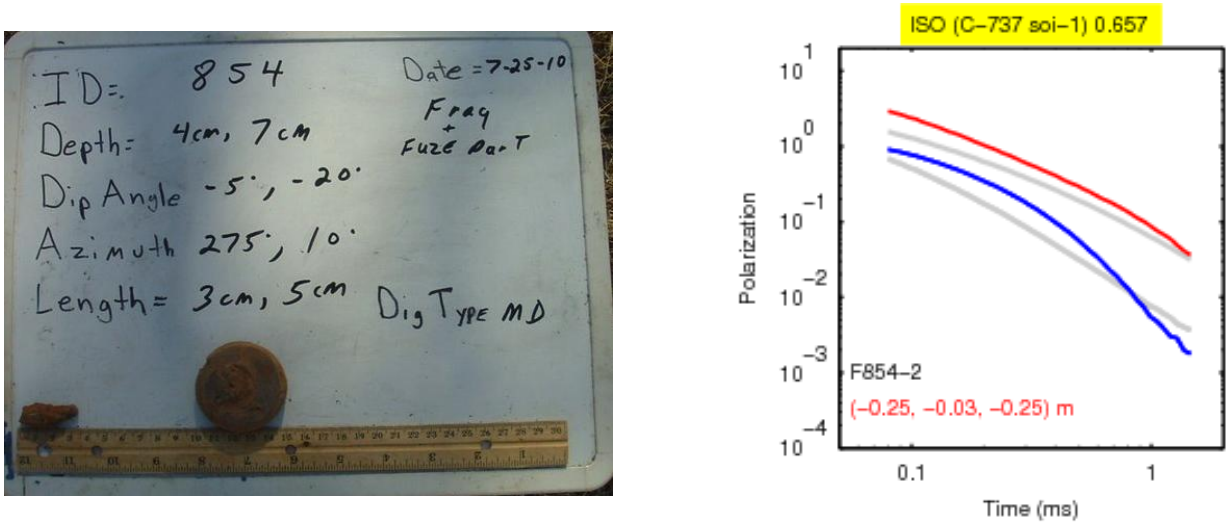


Figure 72: BE-854. The Ground truth and the recovered polarizabilities that are closet to those of ISO was dug at stage 3.

In the end, the total number of dug anomalies was 231 + 36 (training) + 21 (can't analyze) = 288. The total number of TOIs correctly classified is 108 + 8 (training) + 1 (can't analyze): 117.

We dug 288/850 = 33.88 % of the total anomalies. Among them, TOI percentage is 13.76% . ROC curves are presented in Figure 73.

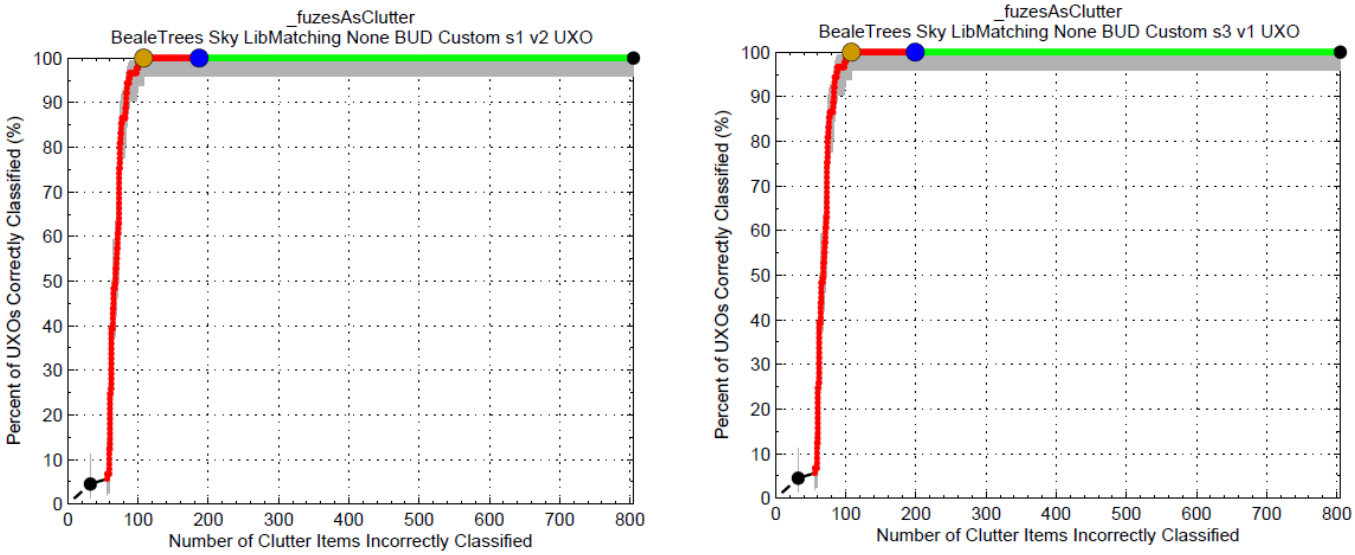


Figure 73: ROC curves at stage 1 and 3 for BUD at Camp Beale.

Figure 75 lists some target anomalies whose recovered polarizabilities partially or almost entirely overlap with a set of reference polarizabilities. To make sure that there were no omissions of any TOI, we choose to dig them at a risk of increasing false positive anomalies.

Small fuzes

There are 6 particular target anomalies: BE-758, BE-196, BE-697, BE-482, BE-805, BE-558 that arise from 5 cm fuze and were treated as clutter according to the program office. In our digging stages, these 6 anomalies were cast into our own clutter partition set. The polarizabilities of these anomalies are closer to the reference fuze polarizabilities than other TOI items in our polarizability library in Figure 75. Still, the match is not nearly good enough to a library item to be moved prior to the stop dig point. It is desirable to explore other complementary ways to identify some of these particular objects.

Retrospective Improvements

Looking back at stage 1, we correctly identified 108 TOIs among 219 anomalies. The set of H-TOI contained the majority (93) TOIs and only one FA anomaly. It was a reasonable partition. However, the remaining 15 TOIs were found from the 4 lower confidence TOI sets at the

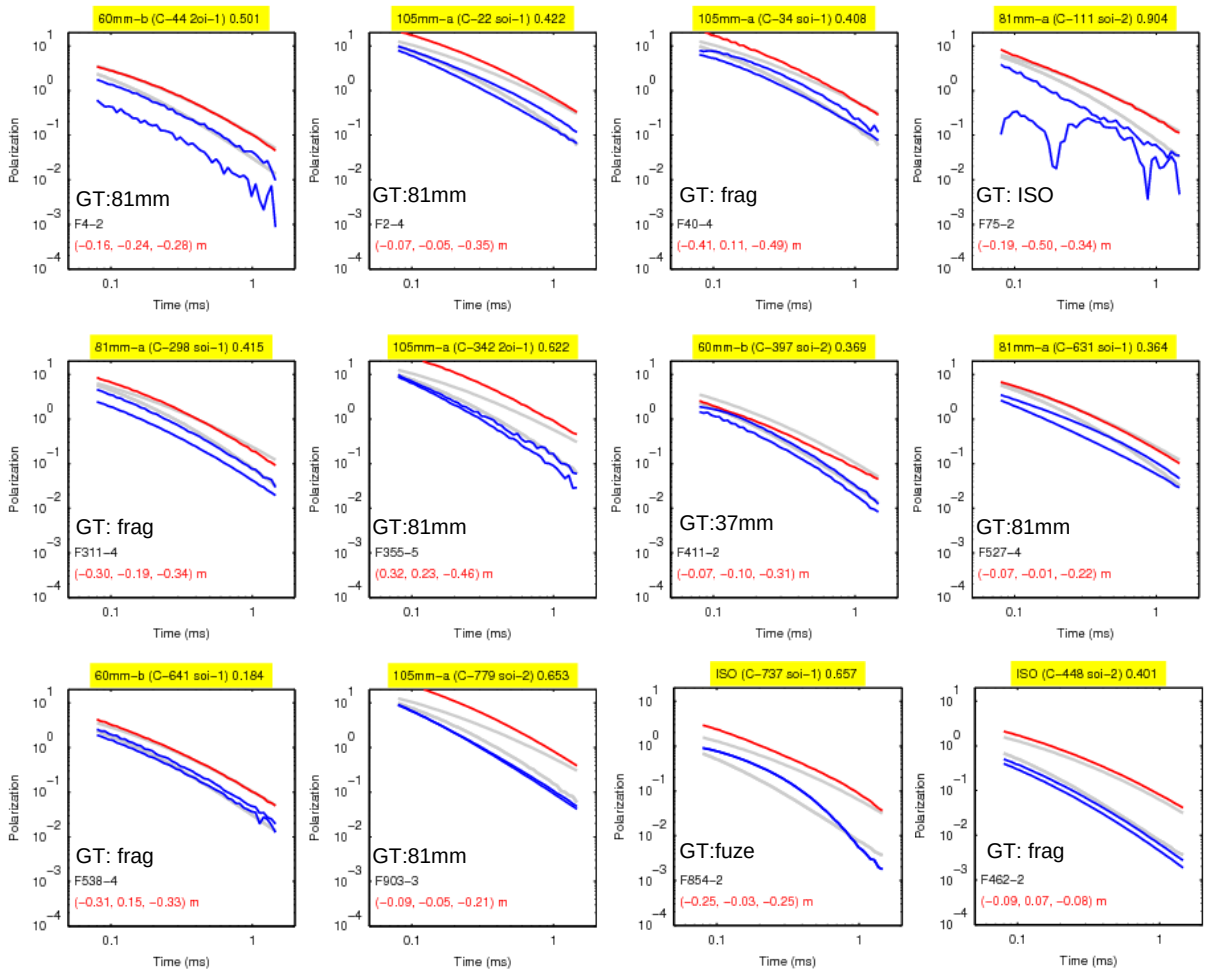


Figure 74: Some examples of the dug anomalies that can cause classification ambiguity. The recovered polarizabilities and the ground truth (GT). The gray curves represent the closest reference polarizations in each plot.

expense of extra digging 110 false positive targets. Our retrospective goal is to identify those 15 TOIs from the 4 sets with the lowest number of FP targets as possible. We attempted such an experiment using two pieces of information from the H-TOI set:

1) Polarizability strength

2) Depths and ranges (relative to the center of sensor) of buried objects.

First, we mapped the size-decay feature of anomalies in the H-TOI set (red solid circles) and the L-TOI sets (orange and green crosses) in Figure 76. The distribution of targets in the H-TOI set all have size > 1 and decay > 0.1. A number of L-TOI targets also fall in this range. We call it TOI size-decay range. These targets in the L-TOI sets would keep their digging status. On the other hand, for targets in the LTOI sets with size < 1, we choose not to dig them in the L-TOI sets. Second, we examined the depth and range of the targets in the H-TOI set and the L-TOI sets. Figure 77 maps this information into six subplots, each of them representing the most likely TOI item grouped according to the mfp values. The targets that are out of TOI size-decay range mainly appear in the two small TOI items of fuzes and 37mm projectiles. Further, we proceed to remove some L-TOI targets in the digging list based upon the H-TOI

set. For each possible item-type, the targets will not be dug if they meet the following depth/range constraints:

- 37mm projectiles: depth > 0.3 m and depth < 0.1 m
- 60mm mortars: range > 0.4 m
- 81mm mortars: no action
- 105mm projectiles: no action
- Fuze: depth > 0.25 m
- ISO: depth > 0.30 m and range > 0.4 m.

Combining these two pieces of information, we exclude 70 targets in the L-TOI sets among 125 potential yet uncertain anomalies. The remaining 55 target anomalies in the L-TOI sets would be dug. Table 2 lists the result of modified stage-1 dig list. PPVs are all improved for the L-TOI sets. The total PPV is increased to 72.48% from the previous 49.31%. Consequently, this modified dig list leads to an extra saving of cleanup cost by 56%. While the above constraints on depth and range produce an improvement via retrospective analysis, this is a site specific finding and further studies are required to determine how applicable these values are for future sites.

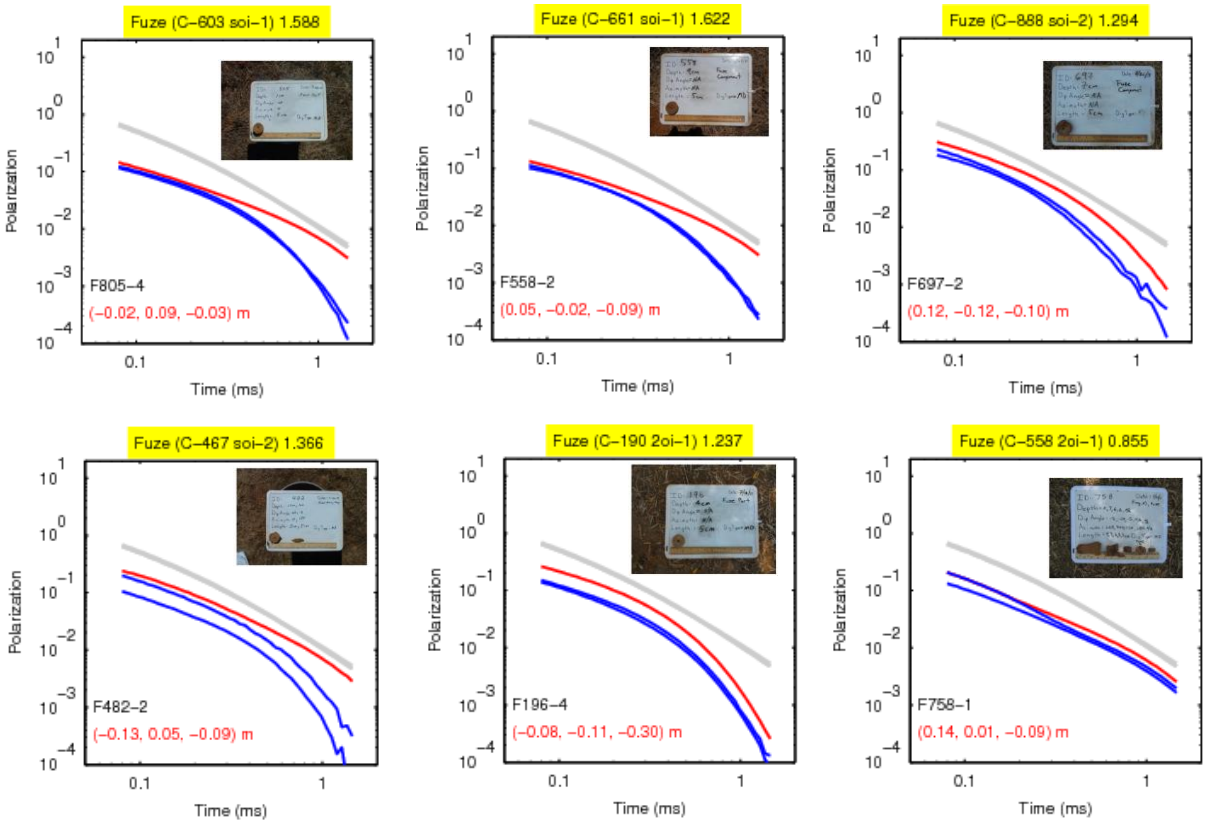


Figure 75: The recovered polarizabilities and the ground truth of the missed small fuze targets. The gray curves represent the reference polarization of fuze in the library.

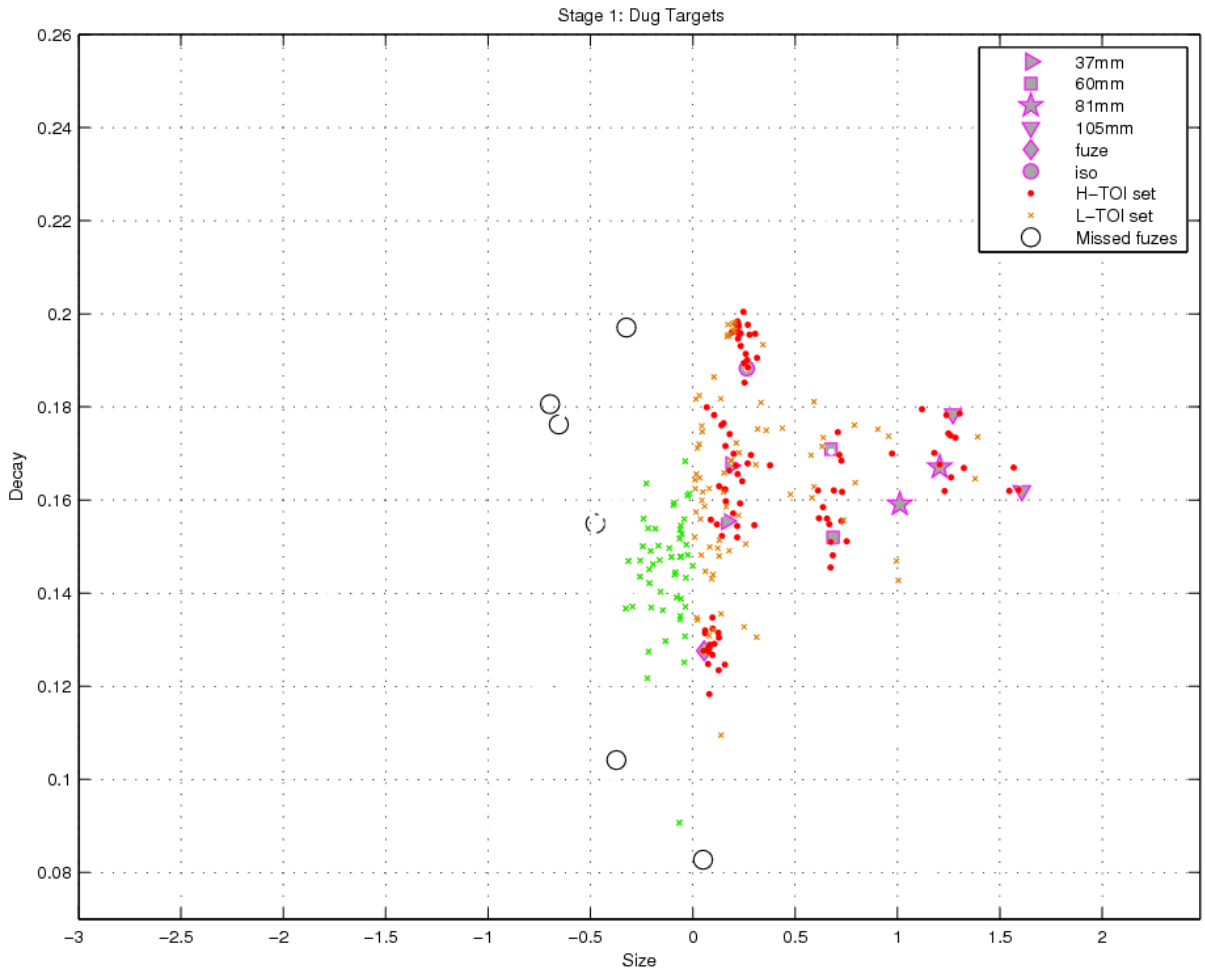


Figure 76: Figure 13. Feature plot of size and decay for BUD data sets at Camp Beale for H-TOI (high TOI) set and L-TOI (likely or low TOI) set at stage-1 dig list. The green crosses denote anomalies in L-TOI with their size feature smaller than 1.

Table 10: Result of modified stage-1 dig list. PPV: Positive predictive value

Set	Dug	TOI	Frag	PPV
H-TOI	94	93	1	98.93%
L-TOI	30	13	17	43.33%
LL-TOI	11	1	20	9.09%
LVL-TOI	9	0	9	0.00%
L-Clutter	5	1	4	20.00%
Total	149	108	41	72.48%

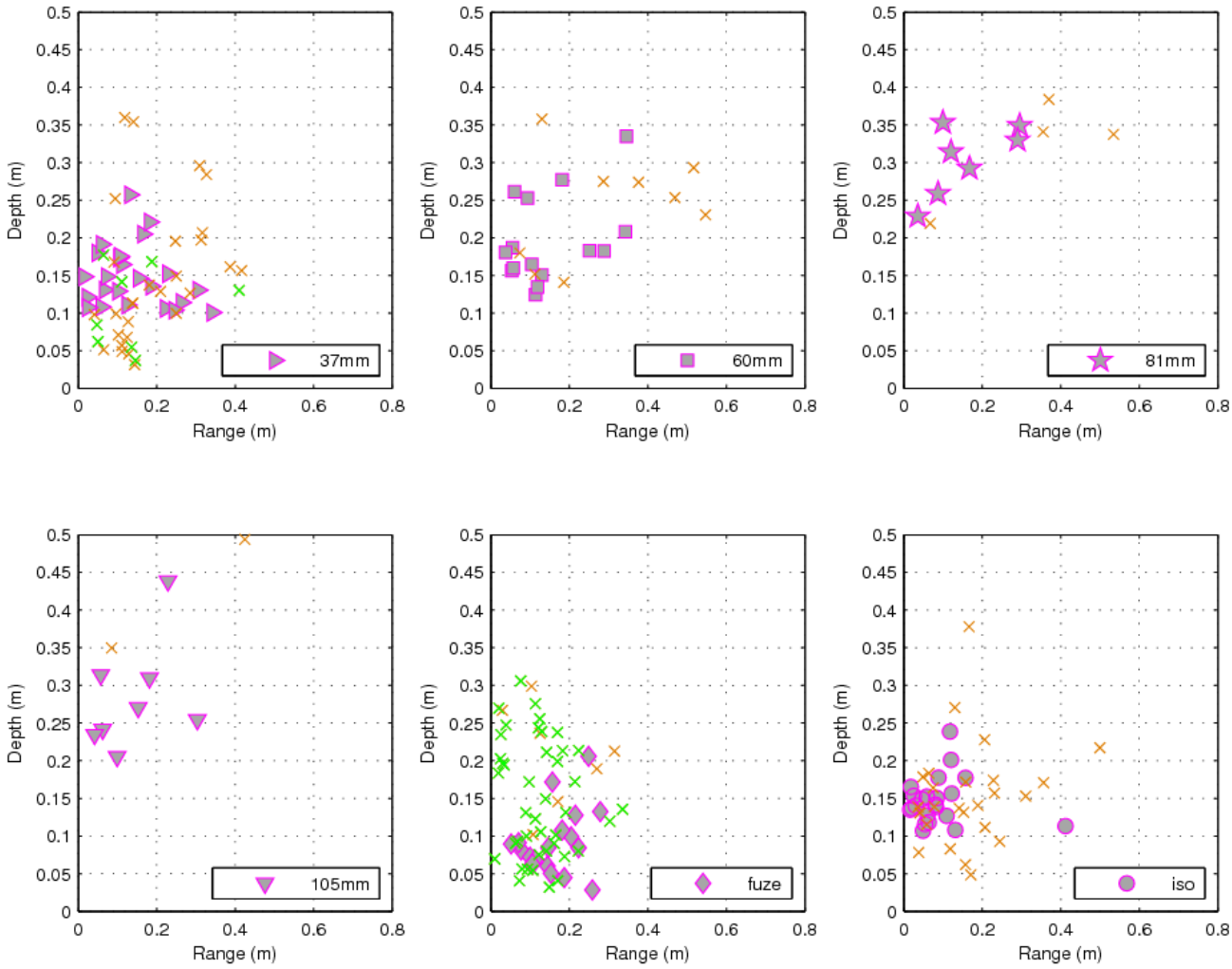


Figure 77: Depth-range plot for 37mm projectiles, 60mm mortars, 81mm mortars, 105mm projectiles, fuzes, and ISO identified in the H-TOI (high TOI) set and the crosses represent the identified items in the L-TOI (likely or low TOI) set at stage-1 dig list. The green crosses denote anomalies in L-TOI with their size feature smaller than 1.

Appendix E: Target Location Error Analysis

The recovered locations of targets were compared to the locations reported in the ground truth files provided by the ESTCP program office.

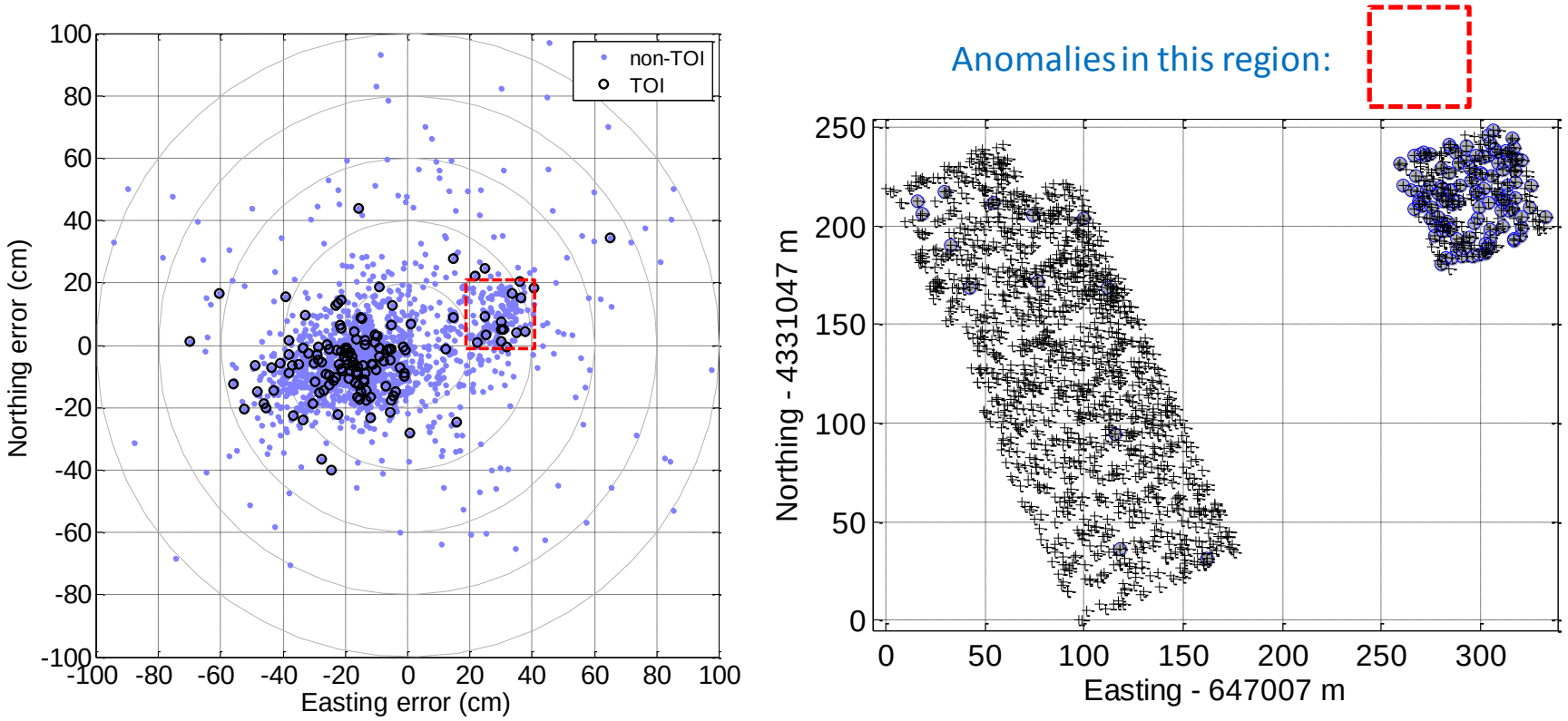
For this analysis, we considered the MetalMapper datasets only. Figure 79 plots the location estimate errors in both Northing and Easting. In Figure 78(a), a scatter plot shows the there is appears to be two clusters of points. There is a main cluster of points centered at an Easting error of approximately -20 cm, and secondary cluster of points centered on an Easting error of approximately +30 cm. The secondary cluster of points is indicated by a red dashed box. In Figure 78(b) all anomalies for the open area is plotted. The square region that is located to the east of the 250 m line (in the local coordinate system of (b)) is shared with the Portable and Open Areas. We see that the secondary cluster in (a) is associated with the shared portion of the survey area. The two regions of the Open area are located on either side of a hill, such that the slope of the western region would have a slope with the opposite sign of the eastern (shared) region. The two cluster of points in the scatter plot of the location errors are consistent with having the reported MetalMapper location being at the GPS antenna location in the exported file. We incorrectly assumed that the reported positions had been corrected to represent the location of the center of the receiver coils.

Although the reported MetalMapper positions represent the location of the GPS antenna, we can still obtain an estimate of the distribution error by calculating statistics of the two survey areas separately. Figure 79 plots histograms of the location errors using all the anomalies. We see in both the Parsons MetalMapper data (top row: (a) and (b)) and the CH2M HILL MetalMapper data (bottom row: (c) and (d)) a bimodal distribution in the Easting error. The distribution of location errors in the western and eastern sections of the Open Area for Parsons MetalMapper data are plotted in Figure 80. Figure 80(a) and (b) plots the same histograms with the anomalies in the western portion of the survey area and Figure 80(c) and (d) plots anomalies in the shared survey area. These plots have a more distinct peak in the histogram. Similar plots were created for the CH2MHILL data inversions (Figure 81). In general the spread of the location error in the Easting direction is greater than the in the Northing direction.

Figure 83 plots the error of the depth estimates for the two MetalMapper data sets. Figure 84 plots the errors for the depth estimates for the three portable sensors. For both the MetalMapper and HandHeld sensors the, the standard deviation of the depth errors are less than the 10 cm success criteria threshold defined in the performance objectives.

Table 11: Summary of target location error for MetalMapper datasets

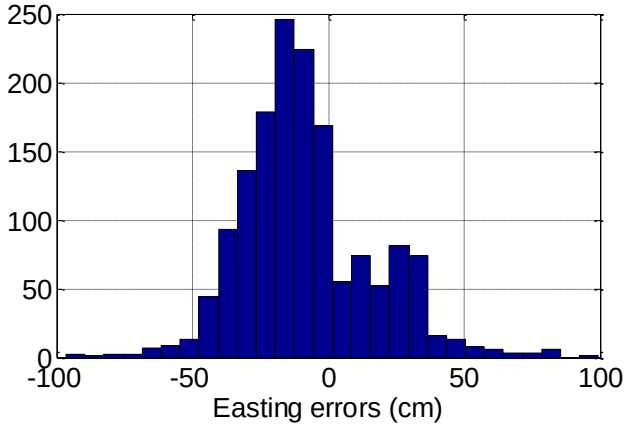
		Easting		Northing	
Data Set	Survey Region	median	σ	median	σ
MetalMapper - Parsons	West	-14.9	16.6	-6.4	13.6
	East (shared with Portable)	24.2	19.4	7.1	13.0
MetalMapper - CH2M Hill	West	-19.2	16.9	-8.7	13.7
	East (shared with Portable)	25.3	17.0	9.6	12.6



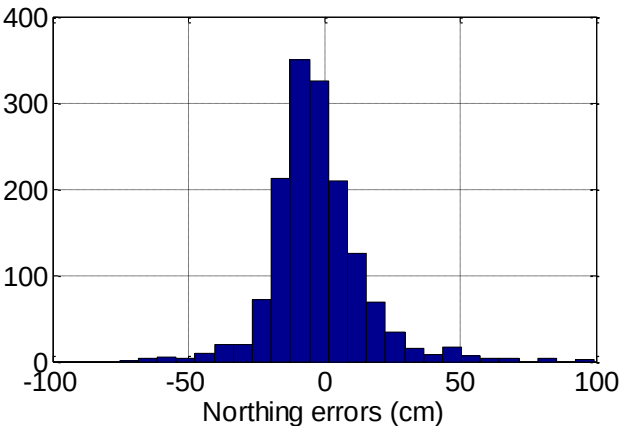
(a) Scatter plot of the location errors

(b) Geographic distribution of anomalies.

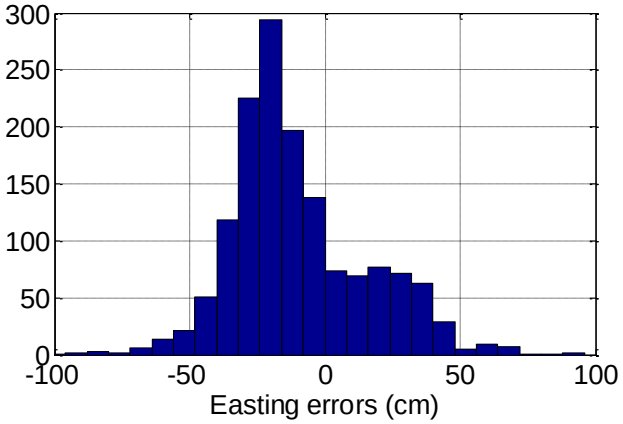
Figure 78: Recovered locations for the Parsons MetalMapper Data set. The scatter plot of location errors in (a) suggest two clusters. In (b) the geographic distribution of anomalies are plotted. The anomalies within the red dashed line in (a) are plotted in (b) with a blue circle.



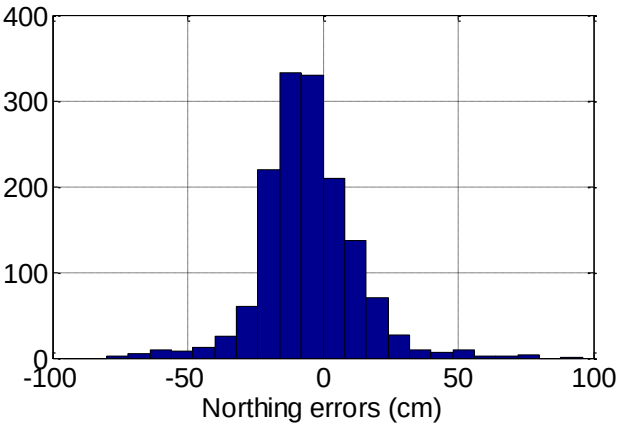
(a) Parsons MetalMapper: Std dev E = 20.9 cm



(b) Parsons MetalMapper Std dev E = 14.2 cm

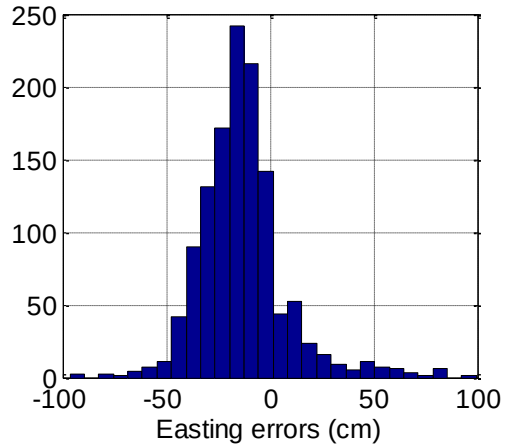


(a) CH2M HILL MetalMapper: Std dev E = 22.3 cm

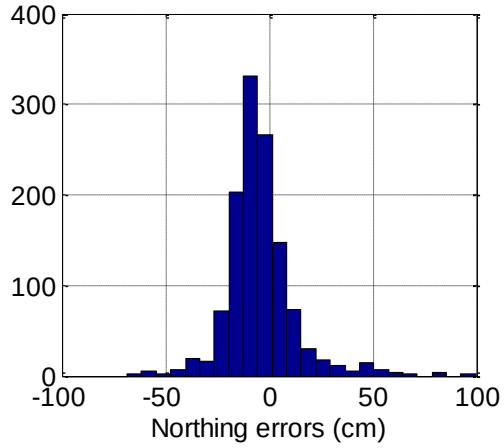


(b) CH2M HILL MetalMapper Std dev N = 14.8 cm

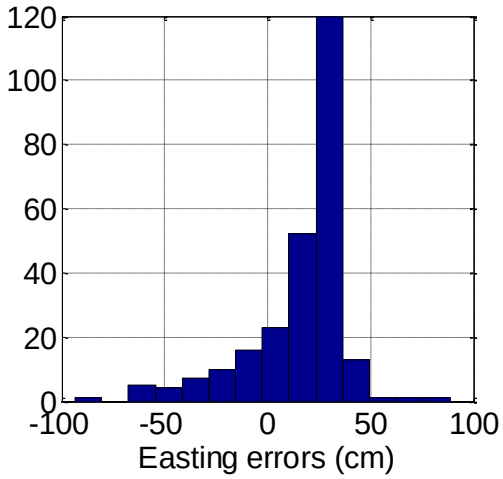
Figure 79: Recovered locations from Parsons ((a)-(b) and CH2M HILL MetalMapper data ((c)-(d)). All Open Area anomalies (including those in the shared area) are plotted.



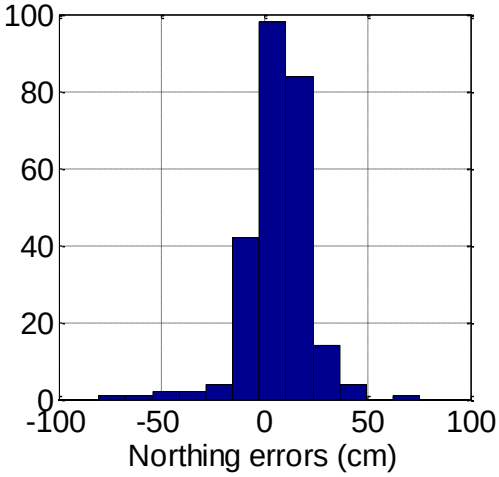
(a) No shared anomalies: Std dev E = 16.7 cm



(b) No Shared anomalies: Std dev N = 13.7 cm

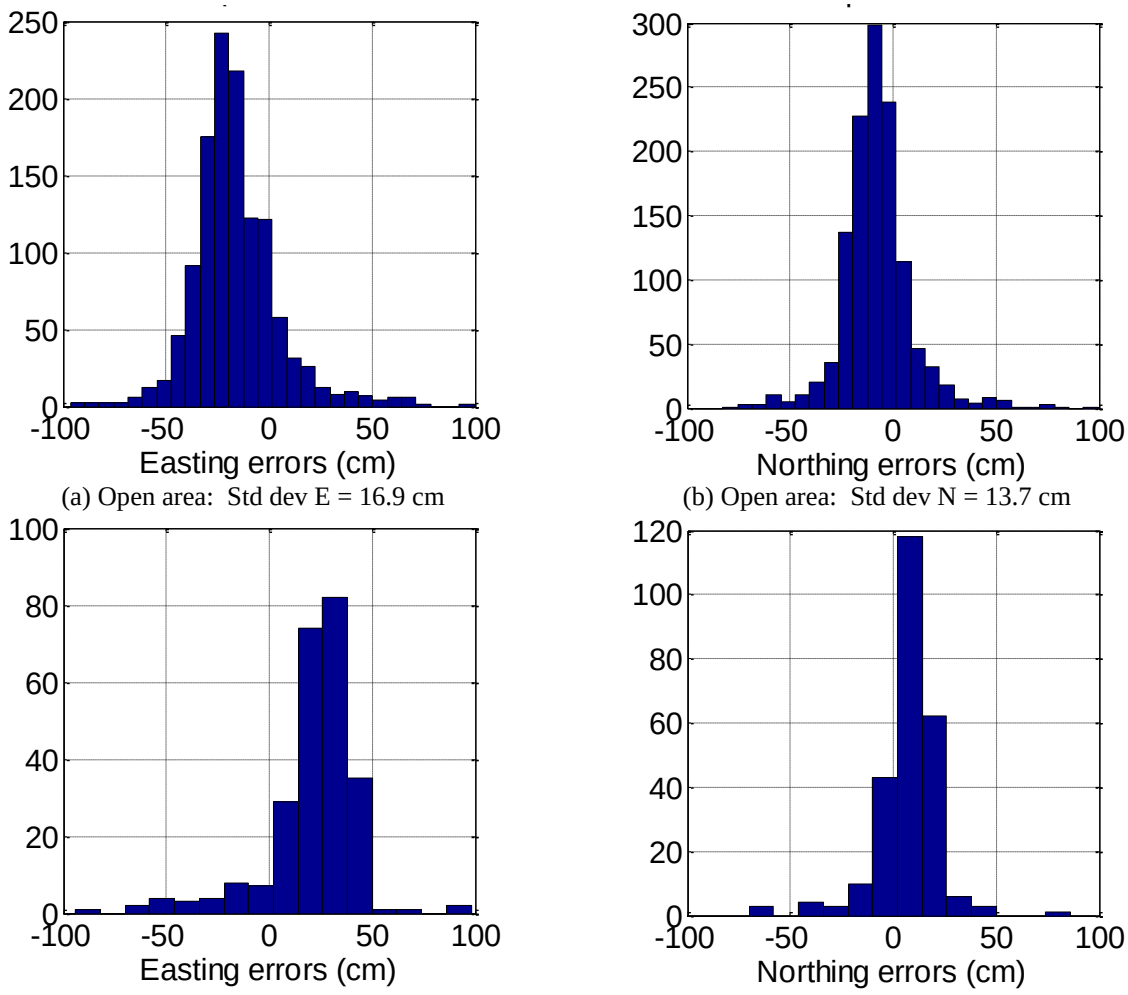


(c) Shared area: Std dev E = 19.4 cm



(d) Shared area: Std dev N = 13.0 cm

Figure 80: Positional errors for the Parsons MetalMapper data. The top row ((a)-(b)) contain anomalies in the western region of the Open Area. The bottom row ((b)-(c)) contain anomalies in the eastern region of the Open Area, which shares anomalies with the Portable Area.



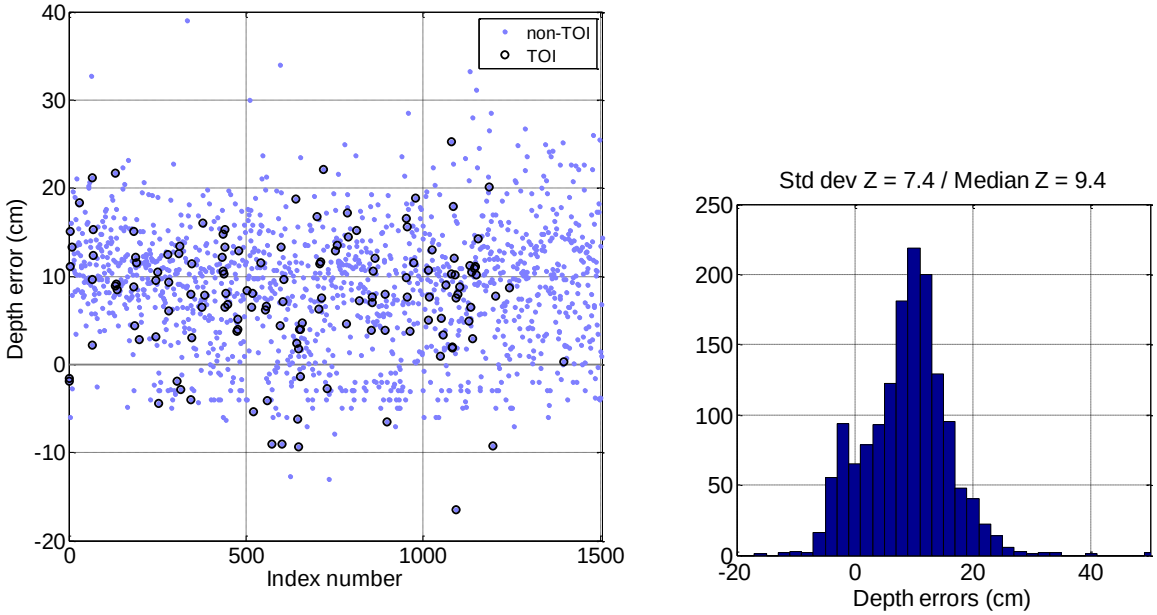
(a) Open area: Std dev E = 16.9 cm

(b) Open area: Std dev N = 13.7 cm

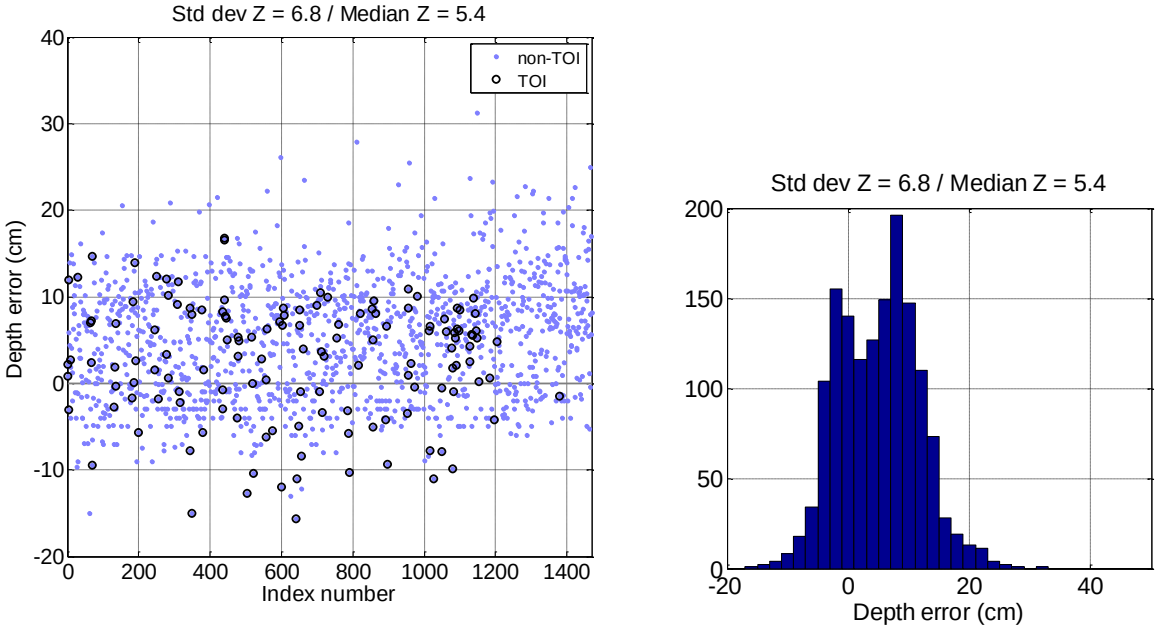
(c) Portable area: Std dev E = 17.0 cm

(d) Portable area: Std dev N = 12.6 cm

Figure 81: Positional errors for the CH2M Hill MetalMapper data. The top row ((a)-(b)) contain anomalies in the western region of the Open Area. The bottom row ((b)-(c)) contain anomalies in the eastern region of the Open Area, which shares anomalies with the Portable Area.

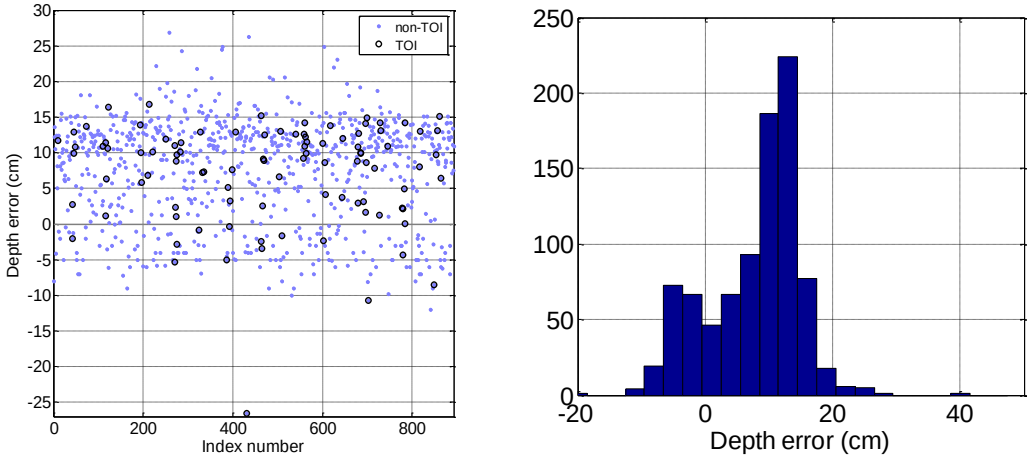


(a) Parsons MetalMapper: Std dev Z = 7.4 cm

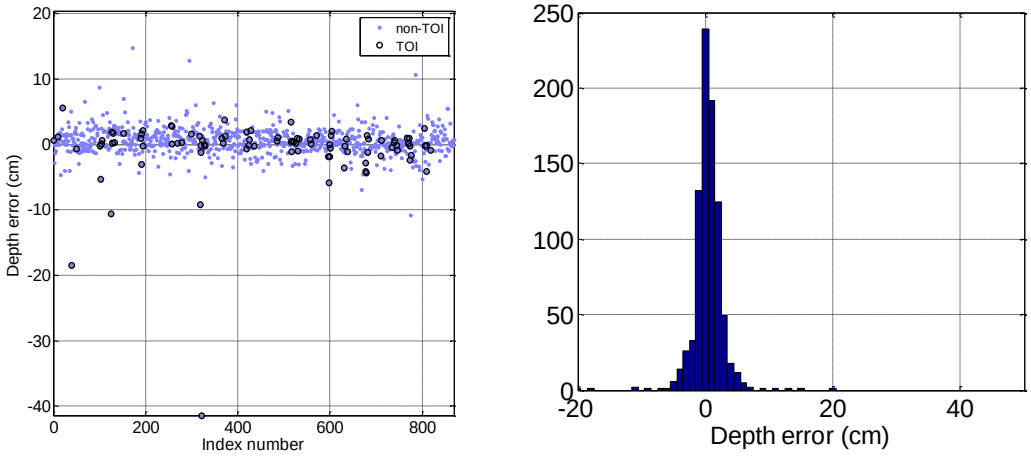


(b) CH2M HILL MetalMapper Std dev Z = 6.8 cm

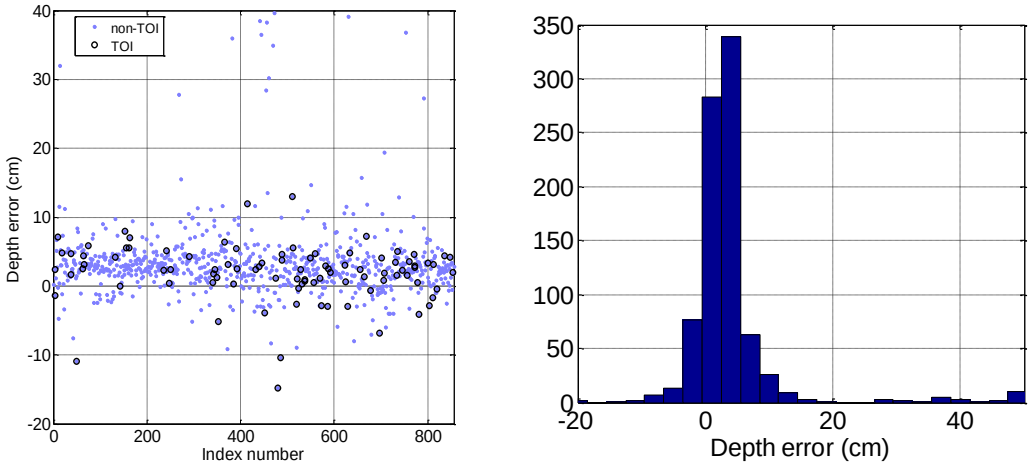
Figure 82: Depth errors for the Parsons MetalMapper data (top row) and CH2MHill data (bottom row).



(a) TEMTADS: Median Error = 9.9 cm. Standard Deviation = 7.2 cm.



(b) BUD Median Error = 0.4 cm. Standard Deviation = 2.7 cm.



(c) MPV Median Error = 2.8 cm. Standard Deviation = 8.8 cm.

Figure 83. Depth analysis for portable sensors.

Appendix F: Points of Contact

POINT OF CONTACT Name	ORGANIZATION Name Address	Phone Fax E-mail	Role in Project
Leonard Pasion	Sky Research Inc, 112A/2386 East Mall Vancouver, BC, V6T-1Z3	541 552 5186 leonard.pasion@skyresearch.com	Principal Investigator (PI)
Kevin Kingdon	Sky Research Inc, 112A/2386 East Mall Vancouver, BC, V6T-1Z3	541 552 5187 kevin.kingdon@skyresearch.com	Project management and personnel coordination
Erik Russell	Sky Research 3 Schoolhouse Lane Etna, NH 03750	541 552 5197 erik.russell@skyresearch.com	Cost tracking